Identification of Therapeutic Targets to Revert Tamoxifen Resistance by Quantitative Proteomic Analysis of Signaling Networks

by

Hideshiro Saito-Benz

Bachelor of Science in Biochemistry
University of California, Los Angeles
2004

Submitted to the Department of Biological Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering
at the
Massachusetts Institute of Technology

April 2009
[June 2009]

Signature of Author:_____
Hideshiro Saito-Benz
Department of Biological Engineering
April 15, 2009

Certified by:_____
Forest M. White
Associate Professor of Biological Engineering
Thesis Supervisor

Accepted by:_____
Steven R. Tannenbaum
Professor of Chemistry, and Biological Engineering
Thesis Committee Member

Accepted by:_____

Peter C. Dedon
Professor of Biological Engineering
Co-Chair, Graduate committee
Associate Head, Department of Biological Engineering

Identification of Therapeutic Targets to Revert Tamoxifen Resistance by Quantitative Proteomic Analysis of Signaling Networks

By

Hideshiro Saito-Benz

Submitted to the Department of Biological Engineering on May 21, 2009
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Biological Engineering

ABSTRACT

Tamoxifen resistance is the biggest problem in endocrine treatment against hormone receptor positive breast cancer patients. HER2 is a membrane receptor tyrosine kinase that is known to correlate with poor disease outcome and unresponsiveness to endocrine treatment. Although much work has been done over the past decades to elucidate pathways involved in HER2 receptor signaling, the map of network-wide signaling events that contributes to the resistance to Tamoxifen treatment has not been characterized, making it difficult to pin-point the downstream drug target to revert the Tamoxifen resistance. To gain a molecular understanding of the mechanisms by which cells gain drug resistance, we have employed a proteomic analysis by mass spectrometry to quantitatively analyze cellular tyrosine phosphorylation signaling events in breast cancer model systems and human tumor samples.

As a result of research, we have identified the major differences in downstream signaling pathways between Tamoxifen sensitive and Tamoxifen resistant breast cancer cell line models. These findings were further analyzed in Tamoxifen sensitive, and Tamoxifen treated/recurred patient samples to study clinical relevance. Specifically, we determined that PI3K/Akt, MEK/ERK, and Src/FAK/Abl pathways are major components of the Tamoxifen resistance. We further showed that they signaling components are possible drug targets to revert Tamoxifen resistance. This study revealed cell-context specific network-wide

changes in signaling events in response to use of therapeutic drugs. This is, to our first knowledge, the first phosphoproteomic analysis of the signaling network in breast cancer to address Tamoxifen resistance. We believe that same approach is applicable to other drug resistance problems in various disease settings.

Thesis Supervisor: Forest M. White

Title: Associate Professor of Biological Engineering

I. INTRODUCTION

The current state of breast cancer treatment,

Phosphoproteomic technologies, and

Phosphorylation signaling networks.

I. 1 The current state of breast cancer treatment

I.1.1 Breast cancer and endocrine therapy

Breast cancer accounts for approximately 30% of all new cancer cases each year in developed countries, with an annual incidence of over 200,000 in the United States alone (Jemal et al., 2006). Approximately 70% of breast cancer patients are candidates for endocrine therapy because of their hormone receptor positive cancers (Anderson et al., 2002). These patients express estrogen receptor (ER), a member of nuclear hormone receptor family. Estrogen signals through ER to stimulate growth of hormone receptor breast cancers. Endocrine therapy targeting estrogen receptor biology plays a role in disease management and significantly improves the clinical outcomes of patients with ER positive breast cancers.

Today, two classes of anti-estrogen agents are considered the standards of care for primary treatment of early- or advanced-stage ER^+ breast cancers. Tamoxifen, a selective estrogen receptor modulator (SERM) which blocks ER

activity within breast tumor cells, has been used for over 30 years and has been the standard choice. Newer generation endocrine therapeutic agents, anastrozole, letrozole, or exemestane, known collectively as the Aromatase inhibitors (AIs), inhibit the production of estrogen through inhibition of the aromatase enzyme pathway in multiple organs including ovaries. Tamoxifen has advantage over AIs in its efficacy in both premenopausal and postmenopausal women, whereas AIs are approved for its use only in postmenopausal women. Since its FDA approval in 1986, Tamoxifen has been tested over decades of clinical studies while AIs have been in clinical use since 1995 when Anastrozole became the first approved AI for breast cancers with disease progression following Tamoxifen therapy. Tamoxifen extends the disease-free survival period of breast cancer patients by approximately 2 years relative to placebo in the adjuvant treatment after surgical removal of breast cancers (Fisher et al., 1996). In addition to the adjuvant treatment, Tamoxifen is approved for the prevention of breast cancer in women who have a high risk of developing breast cancers based on their family history (Fisher et al., 1998). In the adjuvant treatment, several studies have demonstrated that AIs are better tolerated at higher dose and offer improved disease-free survival compared to Tamoxifen

(Coates et al., 2007; Coombes et al., 2007; Forbes et al., 2008). The benefits of AIs have also been demonstrated in the late stage metastatic tumors (Bonneterre et al., 2000; Mouridsen et al., 2003; Nabholtz et al., 2000; Paridaens et al., 2003) and neoadjuvant settings (Eiermann et al., 2001; Ellis and Ma, 2007; Smith and Dowsett, 2003). Although Tamoxifen remains the most popular choice amongst patients, I predict that the aromatase inhibitor treatment will become more common once a long-term safety profile is established and the cost of treatment becomes comparable to Tamoxifen treatment, making AIs widely available around the world.

Despite an initial response to hormone therapy in all treatment settings (primary, adjuvant, or neoadjuvant), breast cancers in many patients progress during therapy. Although it is standard practice in recurrent or resistant ER^+ breast cancer to switch to a different endocrine agent, many questions remain unanswered regarding the appropriate choice of subsequent agents to combat the resistant tumors. In any event, the goals of treatment for patients with metastatic breast cancer are to provide clinical benefits - namely stabilizing the disease state with the aim of prolonging life and delaying disease progression, maintaining quality of life (QOL), and postponing the use of cytotoxic chemotherapy that further impairs QOL.

I.1.2 Mechanisms of estrogen receptor functions

The classic mechanism of estrogen receptor function involves the estrogen binding to ER in the nucleus, thereby promoting association with specific estrogen-responsive elements (EREs) in the promoter region of controlled genes (Jensen et al., 1982). The binding of estrogen to the nuclear ER induces a series of actions leading to activation of the ER. Prior to activation, the ER is stably associated with heat shock proteins in an inactive state. Upon estrogen binding, ER dissociates from the heat shock proteins, undergoes conformational changes, homodimerizes, and undergoes phosphorylation at multiple key serine residues (Osborne and Schiff, 2005). Fully activated ER homodimers binds to EREs, resulting in recruitment of RNA polymerase to initiate gene expression. In addition to classical modes of ER function through EREs, ER can regulate gene expressions of genes without EREs via interaction with other transcription factors. The ER can interact with transcription factors such as the Fos-Jun complex to regulate gene expression at alternative regulatory DNA sequences other than classic ERE

containing genes (Kushner et al., 2000).

In addition to its roles in nucleus, ER is present outside of nucleus and can regulate the activity of membrane bound receptor tyrosine kinases and cytosolic kinases such as EGFR, HER2, IGF1-R, PI3K, and MAPK (Bjornstrom and Sjoberg, 2005; Chung et al., 2002; Kahlert et al., 2000; Levin, 2003; Migliaccio et al., 1996; Schiff et al., 2004; Sun et al., 2001). Besides kinases, ER binds to specific G proteins and activates Src which results in matrix metalloproteinases that cleave transmembrane precursors of heparin binding-EGF (Levin, 2003; Razandi et al., 2003). In additions to the modulations of kinase activities by ER, ER can be activated in the absence of estrogen via phosphorylation by different intracellular kinases. This mode of ER activation is called the ligand-independent activation (Johnston, 2005). The ER is phosphorylated at Serine 118, Serine 167, and Threonine 311, and these phosphorylation events may be due to activation of MAPK, PI3K/Akt, p90RSK pathways and the phosphorylation could occur in response to various cytokines and growth factor stimulation in the absence of estrogen (Campbell et al., 2001; Joel et al., 1998; Kato et al., 1995). The ER phosphorylated in a ligand-independent manner can also translocate to nucleus,

and activate transcription at genes with the EREs.

I.1.3 Current understanding of endocrine resistance mechanisms

With increased understanding of the complex, interconnecting estrogen-ER signaling pathways that regulate cellular responses, it has become clear that tumor cells take different approaches to become resistant to the anti-estrogen therapy. Unraveling the complexity is essential for making the best decisions regarding treatment options, sequencing of currently available therapeutics, and developing future therapeutic compounds. The resistance to therapy is described either as initial (a tumor does not respond to a drug from the onset of therapy, which also often called intrinsic or *de novo*) or acquired (a tumor that initially responded to therapy resumes growth) (Johnston, 1997; Normanno et al., 2005). Approximately 70% of early stage and 50% of advanced stage ER^+ breast cancer patients benefit from Tamoxifen treatment, whereas the remaining patients suffer from initial resistance. In addition, up to 40% of the Tamoxifen responsive breast cancer patients suffer from recurrence due to acquired resistance. Interestingly, the

majority of these patients remain ER^+, indicating that the therapeutic target remains present in the tumor cells. Acquired resistance is currently thought to be a progressive, step-wise phenomenon. The acquired resistance is induced by the selective pressures from the hormonal agents, which drive breast cancer cells from the estrogen-dependent to the endocrine non-responsive phonotype. The finding that ER expression is maintained in the majority of Tamoxifen resistant tumors suggests that the acquired resistance phenotype is caused by complex, multifactorial changes in the ER signaling network rather than a simple, single-gene effect. This also may explain why approximately two thirds of breast cancer patients who developed acquired Tamoxifen resistance and recurrent disease remain responsive to AIs (Lewis and Jordan, 2005), suggesting that the acquired resistance tumor cells remain the estrogen dependent for growth.

I.1.4 Altered growth factor receptor signaling in endocrine resistance.

Altered activities of growth factor receptors resulting in deregulated signaling networks are common to therapeutic resistance in many forms of cancer,

and may contribute to the anti-estrogen resistance in the ER^+ breast tumors. Almost 30% of breast cancers are noted to overexpress the human epidermal growth factor receptor 2 (HER2, also termed *ErbB2, neu*), a member of the epidermal growth factor receptor (EGFR) tyrosine kinase family. Data from preclinical studies and from the retrospective analyses of clinical trials suggest that HER2 is a negative predictor of the response to Tamoxifen treatment (Benz et al., 1992; Dowsett et al., 2008). HER2 and ER cross-talk to stimulate their activities in a cellular model of Tamoxifen resistance (Shou et al., 2004). Our understanding of the mechanisms and outcomes of cross-talk between the ER and HER2 signaling networks remains incomplete. Although the ligand-independent activation of ER has shown to play an important role in ER phosphorylation and activation, how ER communicates back to the HER2 signaling pathway remains unclear.

I.1.5 Effects of the long-term estrogen deprivation by endocrine therapy

Human breast cancer cells that are deprived of estrogen can adapt by developing estrogen hypersensitivity (Masamura et al., 1995; Santen et al., 2005).

Development of estrogen hypersensitivity may explain the clinical observation that estrogen-dependent breast cancer that initially regressed after ovaryectomy-induced estrogen deprivation in premenopausal women re-grew in response to low-dose estrogen replacement therapy and subsequently regressed further after exposure to AIs. Additional data suggest that breast cancer cell lines subjected to long-term estrogen-deprivation demonstrate step-wise progression whereby they first become hypersensitive to low-dose estrogen and eventually become estrogen independent (Chan et al., 2002). One possible mechanism to explain such acquired hypersensitivity and resulting estrogen-independent ER activity is increased ligand-independent ER phosphorylation.

I.1.6 Sequential endocrine therapy to treat resistance

When initial or acquired resistance to the endocrine therapy occurs, no effective treatment guidelines exist regarding subsequent hormone treatments for patients with recurrent or systemic diseases. Although the National Comprehensive Cancer Network recommends a second-line hormone therapy such

as nonsteroidal or steroidal AIs; fulvestrant or other SERMs, the guideline does not recommend a preferred treatment or the appropriate sequencing of agents for recurrent breast cancer after endocrine treatment (Cancer guideline, Accessed March 1st 2009). Forty to fifty percent of patients with breast cancer who have a response to the initial hormone therapy will respond to subsequent treatments with agents with a different mechanisms of action from the initial therapy (Buchholz et al., 1999). Although sequential endocrine therapy is beneficial for postponing aggressive chemotherapy, the eventual development of resistance to sequential second-line endocrine therapy seems unavoidable.

I.1.7 Hormonal treatments in $ER^+/HER2^+$ breast cancer patients

HER2 overexpression in breast cancers strongly correlates with poor disease outcomes (Slamon et al., 1987). From this aspect, it is not surprising that $ER^+/HER2^+$ patients do not respond particularly well to endocrine therapies compared to $ER^+/HER2^-$ patients. The first evidence of a potential negative correlation between the expression of HER2 and the response to Tamoxifen

originated from the Gruppo Universitario Napoletano (GUN)-1 study which evaluated the HER2 expression levels in 145 out of 309 lymph-node negative breast cancer patients. The adjuvant Tamoxifen treatments were associated with the improved disease-free survival (DFS) and the overall survival (OS) in the $HER2^-$ patients but the treatments were associated with worse DFS and OS in the $HER2^+$ patients (Carlomagno et al., 1996; De Placido et al., 2003). In agreement with this finding, studies conducted in Sweden found that the $HER2^+$ patients did not further benefit from five years of Tamoxifen treatment compared to the placebo, whereas $HER2^-$ patients significantly benefited from such treatment. Several retrospective analyses of studies with Tamoxifen treatments in advanced breast cancers showed worse outcomes for the patients expressing high levels of HER2 than the $HER2^-$ patients (Elledge et al., 1998; Wright et al., 1992). These results show that Tamoxifen is not the optimal treatment option for all ER^+ patients and that other disease markers, such as HER2, should be taken into account before the start of endocrine treatments, including but not limited to Tamoxifen treatments.

The availability of novel anti-estrogenic therapeutics such as the AIs has instigated investigations into the efficacy of these therapeutic compounds compared

to Tamoxifen in breast cancer patients with HER2 overexpression. Lipton *et al* showed that the non-steroidal aromatase inhibitor Letrozole is superior to Tamoxifen as a first-line therapy independent of serum HER2 levels (Lipton et al., 2003). Furthermore, the retrospective analyses of data collected in the trials of neoadjuvant endocrine therapies comparing Letrozole treatments versus Tamoxifen treatments found that the response rate is significantly higher for Letrozole compared with Tamoxifen in the subgroup of patients with EGFR or HER2 expression (Ellis et al., 2001). Unfortunately, the relatively small sample numbers from these two studies makes it challenging to generalize that AIs are superior to Tamoxifen for ER^+ patients regardless of HER2 status. However, these studies suggest that the ER functions are affected by the binding of estrogen or anti-estrogen in cell specific manners and the results may significantly influence the clinical guideline.

I.1.8 The current status of the integration of endocrine therapies and kinase inhibitors

Since perturbed growth factor signaling events may play significant roles in the both initial and acquired Tamoxifen resistance in breast cancer cells, the use of combinatorial treatments to inhibit both ER and the signal transduction mediated by kinases represents one of the most promising therapeutic approaches. In this respect, therapeutic compounds inhibiting signaling molecules such as HER2, EGFR, IGF1-R, Ras/Raf/MEK/MAPK, PI3K/Akt, and c-Met are currently in clinical developments. The results from the cell line studies indicated that combinatorial treatment might be effective in both hormone-sensitive and hormone-resistant breast cancer patients. In many cases, the combinatorial treatments are reported to be additive or synergistic in killing the Tamoxifen resistant model cell lines. For example, farnesyl transferase inhibitors that target Ras activity were synergistic with Tamoxifen or aromatase inhibitor treatments against multiple breast cancer cell lines (Ellis et al., 2003; Liu et al., 2007). Similar results were obtained with the combination of mammalian target of rapamycin (mTOR) inhibitors with Letrozole in multiple breast cancer cell line models (Boulay et al., 2005). A synergistic anti-tumor effect has been reported for the anti-HER2 humanized monoclonal antibody Herceptin in combination with Tamoxifen against $ER^+/HER2^+$ cell line

models (Argiris et al., 2004; Ropero et al., 2004). Finally, the combined treatments of the EGFR inhibitor Iressa and Tamoxifen were effective against the breast cancer cell line models with acquired Tamoxifen resistance (Gee et al., 2003). These cell line studies show that the combinations of signal transduction inhibitors with endocrine therapy may be more effective in the first line treatment of ER^+ breast cancer patients than endocrine therapy alone. In addition, such combinations of treatments have the potential to prevent the development of acquired resistance as the signaling molecules proposed to facilitate acquired resistance are inhibited along with ER. However, as discussed in the clinical trials of the combinatorial therapy in the following section, the combinatory treatments still have room for improvements. One of the main reasons for the lack of desired efficacy is that many breast tumors are heterogeneous populations. The combinatorial treatments described above are mostly conducted in a single model cell line system (for example, the overexpression of HER2 in a MCF7 cell line). This limits the identification of better therapeutic approaches that are effective against heterogeneous tumor populations. Therefore, the successful identification of a therapeutic approach to endocrine resistance requires evaluation of perturbed

signal transductions that are common to multiple endocrine resistant cell line models.

I.1.9 Clinical trials of combinatory treatments against ER$^+$ breast cancer patients

There are a few published reports on the clinical trials for the combination of endocrine therapy and kinase inhibitors in the breast cancer patients. Encouraging preliminary results have been disclosed from the trial testing a combinatorial treatment of Herceptin and Letrozole in ER$^+$/HER2$^+$ advanced breast cancer patients (Marcom et al., 2007). Combined Herceptin and Letrozole treatments produced durable responses for a half of the patients, whereas the combination was ineffective in the other half. The finding suggests that ER$^+$/HER2$^+$ advanced breast cancers are heterogeneous and could be divided further into subgroups. The finding also suggests that additional therapeutic agents are required in order to manage ER$^+$/HER2$^+$ advanced breast cancers. In the case of neoadjuvant combination treatment of ER$^+$/EGFR$^+$ patients with Iressa

and the non-steroidal aromatase inhibitor Anastrazole, the combinatorial treatment was effective in reducing the pretreatment values in proliferation-related Ki67 labelling index than Iressa alone (Polychronis et al., 2005). The size of breast tumors was reduced by 30-99% in 14/28 patients in the combinatorial treatment group, while administrations of Iressa alone reduced the size of breast tumor in 12/22 cases, and both the combinatorial treatment and Iressa treatment reduced phosphorylation of ER at Ser 118 at a similar extent. These results indicates that combinatorial treatments of endocrine therapy and kinase inhibitors are tolerated and may be effective in subgroups of $ER^+/EGFR^+$ breast cancer patients.

The following two sections in the chapter 1 are from the published review titled "*Quantitative phosphoproteomics by mass spectrometry: Past, Present, and Future"* in the *Proteomics volume 8, Issue 21, Page 4433-4443.*

I. 2 Phosphoproteomic technologies

I. 3 Phosphorylation signaling networks

Author Contributions:

Aleksandra Nita-Lazar wrote the I.2 Phosphoproteomic Technologies.

Hideshiro Saito-Benz wrote the I.3 Phosphorylation signaling networks.

Forest M. White edited both sections.

REVIEW

Quantitative phosphoproteomics by mass spectrometry: Past, present, and future

Aleksandra Nita-Lazar[1]*, *Hideshiro Saito-Benz*[1]* *and Forest M. White*[1,2]

[1] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[2] Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA

Received: March 12, 2008
Revised: May 2, 2008
Accepted: May 6, 2008

Protein phosphorylation-mediated signaling networks regulate much of the cellular response to external stimuli, and dysregulation in these networks has been linked to multiple disease states. Significant advancements have been made over the past decade to enable the analysis and quantification of cellular protein phosphorylation events, but comprehensive analysis of the phosphoproteome is still lacking, as is the ability to monitor signaling at the network level while comprehending the biological implications of each phosphorylation site. In this review we highlight many of the technological advances over the past decade and describe some of the latest applications of these tools to uncover signaling networks in a variety of biological settings. We finish with a concise discussion of the future of the field, including additional advances that are required to link protein phosphorylation analysis with biological insight.

Keywords:
Electrospray ionization-tandem mass spectrometry / Protein phosphorylation / Signal transduction profiling / Spectrometry

1 Introduction

In the postgenomic era rapid advancement in the characterization of new genes and their protein products has driven an increased demand to functionally classify these proteins. Classical genetic, biochemical, and protein chemical approaches have been historically used to tackle this challenge for selected biomolecules, but these methods tend to be time-consuming, laborious, and usually require large amounts of material. Application of these approaches to characterize thousands of proteins is therefore unrealistic. However, recently developed proteomic methods, quickly improving with technical advancements in equipment, permit large-scale protein analysis while maintaining molecular resolution. While these large-scale methods do not directly provide functional characterization, they can be used to generate hypotheses regarding the function of selected proteins. Follow-on biochemical studies can then be performed on these proteins to validate hypotheses.

Functional classification is further complicated by protein PTMs, which can modify enzymatic activity, binding affinities, and protein conformation. Among PTMs, phosphorylation is perhaps the best studied due to the association between dysregulated phosphorylation and human pathologies [1]. Protein phosphorylation on Ser (~90%), Thr (~10%), and Tyr (~<0.05% of protein phosphorylation) residues is reversible and its dynamic addition can produce fast and precise changes in protein properties, which in turn affect many critical processes, such as protein–protein interactions, cell signaling, cytoskeleton remodeling, cell cycle

Correspondence: Dr. Forest M. White, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail: fwhite@mit.edu
Fax: +1-617-258-0225

Abbreviations: EGFR, epidermal growth factor receptor; **HMEC,** human mammary epithelial cells; **IMAC,** immobilized metal affinity chromatography; **IP,** immunoprecipitation; **NSCLC,** nonsmall cell lung cancer; **PLSR,** partial least squares regressions; **SCX,** strong cation exchange; **SILAC,** stable-isotope labeling of amino acids in cell culture

* Both these authors contributed equally.

events, and cell–cell interactions [2]. Protein phosphorylation analysis is still very challenging, although breakthrough developments over the past decade have now enabled the identification and quantification of thousands of sites from given biological samples. To put these advancements in the field of phosphoproteomics into perspective, Fig. 1 highlights several of the most significant publications over the past 7 years.

Our focus in this review is on quantitative phosphoproteomics by MS. Here we discuss the latest developments in the field, including instrumentation, reagents, and enrichment techniques. Selected applications are highlighted to demonstrate the capabilities of these methods, with an eye toward quantification of signaling networks and use of this information for drug target discovery (for recent reviews, see ref. [3, 4]).

2 Challenges of phosphoproteomics

Phosphoproteomic analysis is plagued by the same challenges facing all proteomic experiments: complexity, dynamic range, and temporal dynamics. The true complexity of the phosphoproteome has yet to be determined, but the Phosphosite database (http://www.phosphosite.org) now lists >30 000 phosphorylation sites on >17 000 proteins, and this number is steadily increasing as each large-scale phosphorylation analysis continues to identify a large number of novel sites. With so many of the proteins in the cell being phosphorylated, the dynamic range of the phosphoproteome is similar to that of the proteome (*i.e.*, $\sim 10^9$), but is further increased by substoichiometric modification. In addition, the temporal dynamics of protein phosphorylation regulate the rapid activation and deactivation of cellular signaling networks, further complicating analysis of the phosphoproteome. So the challenge is not simply to identify and catalog all of the phosphorylation sites, but rather to identify the site, quantify the stoichiometry, and monitor the temporal change in phosphorylation in response to a variety of cellular perturbations. Performing this task on a large number of phosphorylation sites across a broad swath of the signaling network is especially challenging, but is required to understand the mechanisms by which protein phosphorylation controls cell biology.

3 Enrichment methods

Phosphorylated proteins span the gamut of protein expression level, from hundreds of millions to a few copies *per* cell. However, many of the phosphorylation events associated with canonical cellular signaling pathways occur on proteins expressed at relatively low levels. Since phosphorylation of these proteins is often substoichiometric and transient, phosphopeptides obtained from these proteins after proteolytic digest are nearly impossible to detect in the whole cell lysate or tissue sample, which can generate potentially millions of peptides. Selective enrichment of phosphorylated

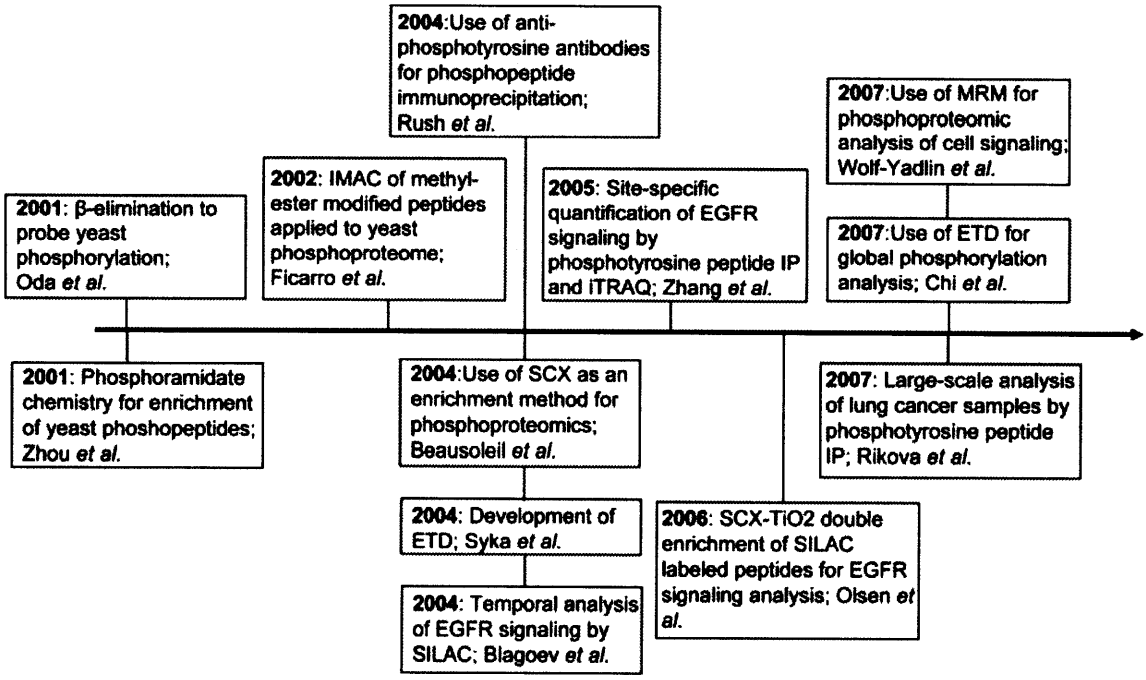


Figure 1. Timeline of selected milestones in quantitative phosphoproteomics during the last decade. Each publication has been selected based on implementation of a new method or application of recently developed methodology to uncover novel aspects of signaling networks.

peptides and proteins is required and has been accomplished in a number of ways, including antiphosphotyrosine antibodies [5], immobilized metal affinity chromatography (IMAC) [6], chemical modification, and strong cation exchange chromatography (SCX) [7].

Immunoprecipitation (IP) of tyrosine phosphorylated proteins and peptides with high affinity antiphosphotyrosine antibodies [8] provides good yield and specificity and has been demonstrated on a broad variety of applications [9–12]. Several reliable antiphosphotyrosine antibodies are sold commercially. These antibodies primarily recognize phosphotyrosine, but each has some bias toward the surrounding amino acids, and therefore performing the IP with multiple antibodies may increase coverage of the tyrosine phosphoproteome. Since the fraction of tyrosine phosphorylated protein to total protein may vary significantly from sample to sample, experimental optimization of conditions, including relative amount of antibody to total sample protein, is often necessary to reduce nonspecific binding while maximizing yield for the particular sample. It is worth noting that while IP has been succesfully implemented for tyrosine phosphorylation, anecdotal evidence indicates that the analogous pan-specific antibodies against phosphoserine and phosphothreonine tend to be of lower affinity, and therefore yield unsatisfactory enrichment for these subsets of phosphorylated peptides. However, recent work by Matsuoka *et al.* [13] has demonstrated the potential of using multiple phospho-specific antibodies recognizing ATM/ATR substrate phosphorylation sites to identify and quantify hundreds of serine and threonine phosphorylation sites matching the ATM/ATR kinase motif. Since many phospho-specific antibodies have off-target affinity, it may be that this strategy could be applied to a variety of serine/threonine kinases, effectively supplementing the need for high affinity pan-specific phospho-serine/threonine antibodies, and enabling network analysis of serine/threonine phosphorylation, one motif at a time.

For many applications, the goal is to generate a global view of serine, threonine, and tyrosine phosphorylation within the sample rather than focusing specifically on a selected subset of phosphorylated peptides. Perhaps the most common technique to enrich for global phosphorylation is IMAC, which is based on the high affinity of phosphate groups for metal ions such as Fe^{3+}, Zn^{2+}, and Ga^{3+}. One of the main limitations associated with IMAC-based phosphopeptide enrichment has been the nonspecific retention of nonphosphorylated acidic peptides, due to the weak affinity between negatively charged carboxylates and positively charged metal ions. However, conversion of carboxylate groups to esters effectively eliminates nonspecific retention of nonphosphorylated peptides on the IMAC column [14]. This method has also been used in an automated platform involving online IMAC, nano-LC, and ESI-MS, enabling reproducible detection and identification of phosphopeptides in a low-femtomole range [15], and may be coupled with a stable-isotope labeling step for relative quantification [14]. Since different metal ions appear to enrich for slightly different subsets of phosphorylated peptides, maximal coverage of the phosphoproteome may be obtained by multiple analyses with different metals, or by mixing multiple metal ions in a single IMAC enrichment step.

Within the past couple of years, titanium dioxide (TiO_2) has emerged as the most common of the metal oxide affinity chromatography (MOAC)-based phosphopeptide enrichment methods. This technique requires significantly shorter preparation time and offers increased capacity relative to IMAC resins with the same bed volume. Since this method exploits the same principle as IMAC, it is similarly prone to nonspecific retention of acidic nonphosphorylated peptides. However, loading peptides in 2,5-dihydroxybenzoic acid has been shown to reduce nonspecific binding to TiO_2, thereby improving phosphopeptide enrichment without chemical modification of the sample [16]. Overall, TiO_2 is often considered to be interchangeable with IMAC, in that similar sample levels (*e.g.*, micrograms of protein) can be analyzed and hundreds of sites *per* sample can be identified when either technique is used as the sole enrichment method, although each method has demonstrated differential bias and selectivity.

As an alternative to metal-ion-based enrichment strategies, SCX has been successfully used to separate phosphorylated peptides from peptide mixtures for subsequent MS analysis [7, 16, 17]. In this technique, binding to the SCX column is dependent on columbic interaction between negatively charged resin and positively charged peptides. If sample loading is performed under strong acidic conditions (pH ~2.7), carboxylates are rendered neutral, while the phosphate group retains a negative charge. As a result, the total charge of phosphorylated tryptic peptides is reduced from + 2 to + 1, and the interaction strength with the SCX resin is correspondingly reduced. Elution with a gradient of increasing salt concentration thus allows phosphopeptides to elute earlier relative to nonphosphorylated peptides, providing semiselective enrichment [7]. To reduce the nonphosphopeptide background, a second, IMAC-based enrichment step has been performed on SCX fractions, enabling the identification of thousands of phosphorylated peptides from given samples [7, 17, 18]. As another variation and improvement of the SCX method, a mixed-bed resin comprised of a blend of anion and cation exchangers (ACE) has been recently proposed for phosphopeptide enrichment, increasing retention of acidic peptides, and reducing retention of basic and neutral peptides by the added anion-exchange resin, which in turn improved the identifications of phosphopeptides by 94% over SCX [19].

Phosphorylation enrichment by SCX-based fractionation, either solely or coupled with other enrichment steps, has successfully been applied to identify large numbers of phosphorylation sites (in the order of thousands). However, it is worth noting that the technique, as implemented to date, requires a large amount of starting material (tens of milligrams of protein) which makes it inapplicable to samples

that are available in small or limited quantity. In addition, SCX fractionation decreases the complexity of the starting samples by dividing it into many fractions, each of which requires a separate MS analysis, leading to the possibility of up to 100 MS analyses for each biological replicate. The sample requirements, analysis time, and labor associated with each biological sample has unfortunately limited the application of this technique such that few studies have incorporated biological replicates.

Several laboratories have taken the approach of chemically modifying the phosphate to provide an affinity enrichment tag. For instance, the phosphate groups on serine and threonine can be removed by β-elimination and replaced by ethanedithiol coupled to a biotin tag, making it possible to purify modified peptides using an avidin affinity column [20]. The primary disadvantage of this approach is that tyrosine phosphorylation does not undergo β-elimination, and therefore these peptides are not enriched by this method. It is also possible to directly attach an affinity tag to the phosphate through phosphoramidate chemistry (PAC). Recent improvements in this approach have improved the yield by reducing the number of steps, making the approach much more user-friendly [21].

Different enrichment methods may yield different pools of phosphopeptides from the same peptide mixture, as recently shown in a comparative study conducted by the Aebersold group, where PAC, IMAC, and two types of TiO_2 methods were employed to isolate phosphopeptides from a tryptic digest of *Drosophila melanogaster* Kc167 cells [22]. Performing multiple analyses with several complementary phosphopeptide enrichment methods may be the best way to maximize depth of coverage, albeit at the cost of increased sample consumption and reduced throughput.

It is often the case that any single enrichment step does not provide sufficient specificity when dealing with complex biological samples. Therefore, double enrichment, as in the above scenario with IMAC and SCX, is often required to improve phosphopeptide analysis. In another example, our laboratory has combined antiphosphotyrosine peptide IP with IMAC to analyze tyrosine phosphorylation in murine adipocytes [23], human Jurkat cells [24], and in the epidermal growth factor receptor (EGFR) signaling network in human mammary epithelial cells (HMECs) [25, 26].

4 Quantification approaches

To date much of the work in phosphoproteomics has focused on developing novel methods for the enrichment of phosphorylated peptides/proteins and subsequent application of these methods to identify large numbers of phosphorylation sites from given biological samples. Data generated in these "cataloging" studies may be informative for laboratories studying selected proteins whose phosphorylation sites appear in the catalog, but it is often difficult to link the information in these large-scale datasets to cellular signaling networks. In order to identify phosphorylation events that may be regulating biological response to cellular perturbation, quantification of phosphorylation pre- and post-cell stimulation is necessary. Several MS-based quantification methods have been implemented for phosphoproteomics, including stable-isotope labeling through chemical modification of peptides, stable-isotope labeling of amino acids in cell culture (SILAC), and label-free methods.

Multiple chemical modification protocols have been utilized to incorporate stable isotopes; among these methods, iTRAQ has become the most commonly used option due to its multiplex capability. The iTRAQ reagent consists of four isobaric isoforms which react with primary amines, thereby enabling quantitative comparison of four protein samples in parallel. Since the labels are isobaric, quantification is performed in MS/MS mode by comparing peak areas of the marker ions resulting from fragmentation of the iTRAQ label, so the same spectrum is used for quantification and sequence identification of the phosphopeptide. As demonstrated in Fig. 2, when coupled to phosphotyrosine peptide IP and IMAC, iTRAQ has been successfully applied for the quantification of phosphorylation states of differentially stimulated Jurkat cells [24], adipocytes [23], and for analysis of the temporal dynamics of the ErbB signaling network [25, 26]. Recently, the eight-plex version of iTRAQ has been developed and applied to proteome analysis [27], demonstrating the potential to further increase throughput in quantitative proteomic analysis, but this reagent has yet to be used for the quantification of protein phosphophorylation.

For metabolic isotope labeling, cells are cultured in a medium where the natural form of an amino acid (typically arginine or lysine) is replaced with a stable isotope form, such that proteins expressed by the cell incorporate the heavier version of this amino acid and therefore alter their molecular mass (see ref. [28] for the detailed, updated review of the method). This technique is generally referred to as SILAC, and enables comparison of up to three samples in a single analysis. Initially, SILAC was developed for mammalian cells [29], but its use has been broadened to bacteria [30] and yeast [17]. There have also been reports of *in vivo* metabolic labeling in whole organisms (*Plasmodium falciparum* [31], plants [32], *D. melanogaster, Caenorhabditis elegans* [33], and rats [34]), but they require feeding labeled reagents to model organisms, which makes multiple experiments costineffective.

There are advantages and disadvantages to both metabolic and chemical modification-based labeling methods. For SILAC, cells need to undergo multiple cell divisions in medium containing stable isotope-labeled amino acids to ensure sufficient isotope incorporation for reliable comparison between cell states. For this reason, it is not practical to apply SILAC to generate quantitative data from primary cells or to compare tumor tissue specimens directly. Moreover, culture conditions need to be carefully monitored to prevent interconversion between arginine and proline, which could negatively affect quantification accuracy. However, since cells

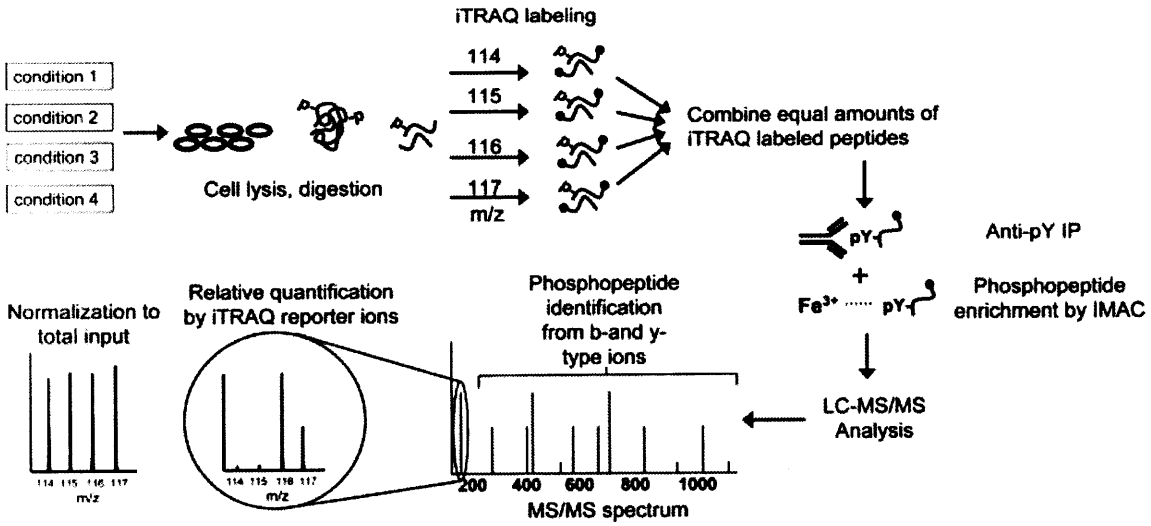


Figure 2. Example workflow of the quantitative phosphotyrosine analysis experiment using iTRAQ labeling. Proteins are extracted from four biological samples (cell lines, stimulation time points, tissue samples), modified and digested. For quantification, the resulting peptides are labeled with iTRAQ reagent and combined. The mixture is then subjected to two steps of enrichment for tyrosine phosphorylated peptides: IP with antiphosphotyrosine antibodies and IMAC. Phosphorylated peptides eluted from the IMAC column are analyzed by LC-MS/MS, typically on a quadrupole TOF mass spectrometer. Each MS/MS spectrum contains both sequence-specific fragmentation events for identification of the peptide and phosphorylation site as well as the low-mass iTRAQ marker ions to quantify phosphorylation across the four samples. iTRAQ quantification is normalized to the supernatant to eliminate variation resulting from the preparation process.

can be mixed prior to cell lysis and sample processing, quantification error associated with differences in these steps can be avoided, potentially leading to higher accuracy. By comparison, postextraction methods permit quantitative analysis of a broader variety of samples, including animal tissues and human tumors, providing the opportunity to follow *in vivo* changes between healthy and diseased states, which in turn can lead to the discovery of new drug targets. Since labeling typically occurs following enzymatic digestion, sample handling needs to be carefully controlled to minimize variation introduced prior to mixing differentially labeled samples.

For many applications, quantification relative to an arbitrary state is not sufficient, and absolute quantification is desired. Typically, absolute quantification would have required the chemical synthesis of heavy-isotope coded peptides [35] to be added to the sample as internal standards. Recently, however, such peptides can be biologically expressed using the method (named QconCAT) introduced by Pratt *et al.* [36, 37], in which *Escherichia coli* is transfected with a modified gene containing the peptide of interest. The transfected *E. coli* is cultured in a medium containing heavy lysine and arginine and the protein is digested following purification, yielding the desired peptide, which can then be added to the sample. A recent review by Mirzaei *et al.* [38] summarizes the current techniques for production of isotope-labeled peptide standards.

Label-free quantification may be employed as a less expensive alternative. These analyses are typically performed either through direct comparison of two samples analyzed on the same platform, or by spiking the sample with standard peptides and quantifying in reference to these standards [39–41]. Unfortunately, label-free quantification does not provide the multiplex advantage associated with SILAC or chemical modification (*e.g.*, iTRAQ), and therefore requires a separate MS analysis for each sample. Moreover, label-free analysis tends to have greater quantification error compared to analysis of stable isotope-labeled samples, due to inconsistent sample processing and chromatography across multiple analyses.

In the near future, MS-based proteomics should be increasingly focused on absolute quantification of protein expression level and stoichiometry of major PTMs. This information will make it feasible to directly compare data between experiments, conditions, and laboratories. Moreover, absolute quantification will enable the development of more complex, kinetic computational models describing the biological systems in much greater detail.

5 Instrument choice

As with all MS, optimal instrumentation for phosphoproteomic analysis is defined by the application. For instance, in the case of global phosphoproteomics, instrument choice may be influenced by the facile neutral loss of phosphoric acid (98 Da) from serine and threonine phosphorylation, often resulting in uninterpretable MS/MS spectra. To circumvent this problem, with a quadrupole ion trap (IT) it is possible to perform MS^3 on the neutral loss peak from the

MS/MS spectrum; this strategy has been successfully implemented for several large-scale phosphoproteomic studies [42]. The facile neutral loss problem may also be addressed by using a quadrupole TOF instrument, as the intensity of the neutral loss peak is diminished by multiple collisions in a high-pressure quadrupole, yielding an increase in sequence-specific fragmentation and improved phosphopeptide identification. More recently, electron capture dissociation (ECD) [43] and, especially, electron transfer dissociation (ETD) [44] have been demonstrated to be particularly useful for the analysis of labile PTMs including serine and threonine phosphorylation, providing good sequence coverage even for large peptides and proteins.

Compared to serine or threonine phosphorylation, the phosphate attached to tyrosine is relatively stable and usually does not produce a neutral loss in MS/MS mode, although loss of 80 Da may be seen from some tyrosine phosphorylated peptides. In fact, MS/MS spectra of tyrosine phosphorylated peptides tend to resemble nonphosphorylated peptides, although fragmentation N- and C-terminal to phosphotyrosine typically produces a characteristic immonium ion of m/z 216.0426. Since this fragment ion is specific to tyrosine phosphorylation, precursor ion scanning coupled with subsequent MS/MS analysis of the selected precursor ions has been used to identify these pTyr-containing peptides from complex mixtures.

Instrument choice is further affected by the chosen quantification method. SILAC experiments require high resolution and high mass accuracy because the number of species in the MS spectra are doubled (or tripled), leading to increased complexity and increased frequency of overlapping peaks. Although many of the initial SILAC experiments were analyzed on a quadrupole TOF instrument, most of these studies are now conducted on instruments with quadrupole IT fragmentation and Fourier-transform based detection (*e.g.*, LTQ-FTMS or LTQ-Orbitrap) due to increased MS/MS acquisition speed in the IT and increased MS acquisition mass accuracy and resolution in the FT mass analyzer. Choice of instrumentation for iTRAQ-based quantification is more restrictive due to the low m/z ratio of the iTRAQ marker ions. Quadrupole IT instruments have traditionally not performed well for these experiments because fragmentation in a quadrupole IT is typically performed at a Q-value of 0.25–0.3, leading to loss of the low mass region of the MS/MS spectrum. The hybrid quadrupole TOF mass spectrometer has been the instrument of choice for iTRAQ-based quantitative phosphoproteomic analyses due to the high resolution and mass accuracy (low ppm range) in both MS and MS/MS mode, providing accurate detection of the charge state and unambiguous assignment of the monoisotopic mass. Importantly, high-resolution MS/MS spectra obtained on this instrument have improved quantification accuracy by separating iTRAQ marker ions from contaminant ions at the same nominal m/z ratio [45]. The recent development of C-Trap-based fragmentation on the LTQ-Orbitrap now enables triple quadrupole-like fragmentation and high resolution, high mass accuracy detection in the orbitrap mass analyzer, potentially providing a viable alternative to quadrupole TOF instruments for iTRAQ-based quantification of phosphorylated peptides [46].

Given the large array of available enrichment and quantification techniques and the possible combinations of these approaches with various types of MS, it is worth reviewing how these options have been implemented to interrogate the phosphoproteome.

6 Phosphoproteomics in the EGFR network

In canonical growth factor signaling, stimulation of cell surface receptors first triggers activation of the receptor and subsequently transmits the signal to a large number of intracellular effecter molecules. The EGFR network is one of the most extensively studied areas of signal transduction, and the one which best exemplifies oncogenic aberrations in cellular signaling [47]. EGFR is a member of ErbB family of RTKs which comprises four receptors (EGFR, HER2, HER3, and HER4) and 13 polypeptide ligands, each of which contains a conserved epidermal growth factor (EGF) domain. This complex signaling network has been one of the primary targets for phosphoproteomic analysis. In fact, one of the first studies to address quantitative dynamics of phosphorylation at the network level was performed in the EGFR system. In this study, a monoclonal antiphosphotyrosine antibody was used to immunoprecipitate SILAC-labeled tyrosine phosphorylated proteins and their binding partners. These proteins were enzymatically digested to peptides and analyzed by LC-MS/MS, resulting in the identification of ~80 signaling proteins, including many known EGFR substrates and several novel effectors [9]. The relative intensity of SILAC-labeled peptides was used to quantify temporal dynamics within the network following EGF stimulation of HeLa cells. However, since enrichment for tyrosine phosphorylation was performed at the protein level and enzymatic digestion produces a broad variety of tryptic peptides, most of which represent nonphosphorylated sections of the immunoprecipitated proteins, very few phosphorylation sites were identified in this study, and therefore much of the key signaling information is missing. In fact, since phosphorylation often happens at multiple tyrosine residues within a single protein, and different phosphorylation sites on a single protein are often differentially regulated with individual functions, quantification of each phosphorylation site in the global signaling network is critically important.

To address the need for site-specific quantification, Zhang *et al.* [26] performed time-resolved temporal analysis of EGFR signaling network by quantitative MS using iTRAQ. In this study, proteins from whole cell lysate were proteolyzed to peptides and labeled with iTRAQ prior to mixing. Tyrosine phosphorylated tryptic peptides were then enriched, first by IP with an antiphosphotyrosine antibody, and

then by IMAC to eliminate nonspecifically retained nonphosphorylated peptides. As a result, 104 tyrosine phosphorylation sites from 76 proteins were identified with temporal phosphorylation profiles at four time points of EGF stimulation. Site specific monitoring of protein phosphorylation in this study provided explicit detail regarding the regulation of proteins within the signaling network, including differential regulation of multiple sites on given proteins, and identification of phosphorylation "modules", clusters of sites with selfsimilar temporal profiles.

Peptide IP has now been successfully implemented in a variety of phosphoproteomic studies, including a recent large scale analysis to identify phosphotyrosine signaling networks in lung cancer cell lines and tumors [48]. In this study, oncogenic tyrosine kinase signaling was characterized by analysis of tyrosine phosphorylation in 41 nonsmall cell lung cancer (NSCLC) cell lines and over 150 NSCLC tumors, resulting in the identification of a total of 4551 sites of tyrosine phosphorylation on greater than 2700 different proteins. Bioinformatic analysis of the dataset identified a subset of NSCLC tumors and cell lines exhibiting high tyrosine phosphorylation, possibly due to the presence of abnormally activated or overexpressed tyrosine kinases. Potential "driver" tyrosine kinases were identified by a ranking process to identify unusually high levels of tyrosine kinase activity in a subgroup of patients. Among the 18 tumors with highest EGFR rank, nine tumors were confirmed to have an activating mutation in the kinase domain. Based on this success, a similar approach was used to identify other candidate driver tyrosine kinases in the remaining tumors, including the fusion tyrosine kinase EML4-ALK, as also recently reported by Soda *et al.* [49]. This example demonstrates that MS-based phosphoproteomic discovery capabilities are highly complementary to the genomic cDNA screening technology that was used to originally identify this transforming fusion tyrosine kinase.

7 Global phosphoproteomics

With advances in MS and phosphopeptide separation methodologies, the scale of global phosphoproteomic studies has increased significantly since the first large-scale global analysis of the yeast phosphoproteome [14] only 5 years ago. For instance, Olsen *et al.* [50] recently quantified global phosphorylation changes in EGF stimulated HeLa cells by combining SILAC for relative quantitation with SCX and TiO_2 chromatography for phosphopeptide enrichment. Enriched phosphopeptides were analyzed by more than 100 LC-MS/MS runs to identify over 6600 phosphorylation sites on 2240 proteins with their annotated subcellular (nuclear *vs.* cytosolic) localization. This information represents the largest global phosphorylation dataset available to date for the EGFR signaling network, and covers a broad spectrum of phosphorylation events, from EGFR autophosphorylation to phosphorylation of terminal effector molecules such as transcription factors. However, even this dataset is still far from comprehensive, as many well characterized phosphorylation sites were not identified in this analysis. For instance, the low number of pTyr sites (103) in the dataset is likely due to the large dynamic range associated with simultaneous analysis of serine, threonine, and tyrosine phosphorylation. Given the large number of previously uncharacterized phosphorylation sites and the massive size of this dataset, extraction of biological hypotheses is not trivial. However, it is likely that improved functional characterization of the EGFR signaling network may arise from linking this dataset to other complementary datasets [51] (*e.g.*, Friedman and Perrimon's work RNAi-based screening to identify components in the RTK–Erk network [52]).

8 Reproducibility of phosphoproteomics

One of the principal limitations with each of the above studies has been the irreproducibility of MS-based data, such that replicate analyses of the same sample (or analysis of biological replicates) will typically identify only 60–70% of the same phosphorylation sites. Much of this irreproducibility stems from operating the mass spectrometer in a nonbiased "discovery" mode, in which the instrument continuously repeats a cycle consisting of a full-scan mass spectrum, followed by fragmentation of a certain number of the most abundant peaks for peptide and phosphorylation site identification. This mode enables identification of novel phosphorylation sites, but the semiautomated peak selection process is inherently irreproducible (see, for example, a study of peptide/protein identification reproducibility by MS [53]), making it difficult to directly compare multiple datasets. Recently we have developed an approach combining "discovery" mode analysis of selected biological samples with high reproducibility multiple reaction monitoring (MRM)-based "monitoring" mode for quantification of hundreds of selected phosphorylated peptides [45]. This method was applied to investigate the temporal dynamics of 226 phosphorylation sites at seven time points of EGF stimulation of HMECs. Because preselected phosphopeptides are specifically monitored in MRM mode, the number of peptides reproducibly identified from four replicates increased from 34% in discovery mode to 88% in monitoring mode. This combined method should be applicable to a variety of biological systems, and will enable reproducible network-wide quantification of cell perturbation effects across a broad variety of stimulation conditions.

9 Phosphoproteomics and systems biology

Computational and systems biology approaches have become increasingly important in the analysis of phosphoproteomic data. To provide higher meaning to the data, quantification of both the phosphorylation network and the

corresponding biological response must be collected. Bioinformatics and mathematical modeling can then be applied to build hypotheses connecting phosphorylation information to cellular phenotypes. In the past, identification of key elements in signaling networks has largely been accomplished in a subjective way through the manual comparison of fold – change phosphorylation and cell behavior. Recently, mathematical modeling methods such as partial least squares regressions (PLSR) have been implemented to objectively correlate phosphoproteomic data with cellular response to stimulation. For instance, Wolf-Yadlin [25] *et al.* and Kumar *et al.* [54] applied PLSR to the quantitative MS data describing the effects of HER2 overexpression on phosphotyrosine signaling in HMECs stimulated by EGF or heregulin (HRG). Cell migration and proliferation were collected under the same conditions and PLSR was used to integrate the data types. The final model described a set of signaling molecules that are most relevant for the changes in migration induced by HER2 overexpression. This type of modeling can provide insight into the functionality of unknown proteins, which can be further tested by biological experiments.

As described above, the value of phosphoproteomic datasets significantly increases when these data are used to generate hypotheses as to the function of selected phosphorylation sites, and even more when these hypotheses are experimentally validated. For instance, Kratchmarova *et al.* [55] interrogated tyrosine-phosphorylation mediated signaling networks following EGF or PDGF stimulation of mesenchymal stem cells. Interestingly, although most of the network responded similarly to these two stimulations, activation of the PI3K pathway was exclusive to PDGF stimulation. Since EGF stimulation of these cells drives osteoblast differentiation, the authors hypothesized that PDGFR-associated PI3K activation could be key to controlling biological response to differential growth factor stimulation. Indeed, this hypothesis was validated by small molecule inhibition of PI3K followed by PDGF stimulation to drive ostoblast differentiation.

Another example of biological validation of MS-based phosphoproteomic data was provided recently by Huang *et al.* [56] in the quantitative phosphoproteomic analysis of the EGFRvIII signaling pathways in U87MG glioblastoma cell lines. Clustering of phosphorylation data identified previously unknown crosstalk between EGFRvIII and c-Met, a receptor tyrosine kinase that is well known to drive malignancy in various cancers. Since EGFRvIII and c-Met may signal cooperatively to drive tumor growth, U87GM cells expressing EGFRvIII were treated with the EGFR kinase inhibitor AG1478 and the c-Met inhibitor PHA665752. Either compound alone had minimal cytotoxic effect, but the combination of the two compounds significantly increased cytotoxicity at lower doses, indicating that EGFRvIII utilizes other receptor tyrosine kinases to potentiate oncogenic signaling. This finding has recently been corroborated through the analysis of glioblastoma cell lines and tumors using antiphosphotyrosine antibody arrays [57].

10 Phosphoproteomics and drug development

As described above, quantitative MS-based phosphoproteomics has been applied to identify oncogenic kinases which may serve as potential drug targets. To validate this hypothesis, cells are often treated with selected kinase inhibitors with the goal of altering cellular phenotype, but it is often difficult to establish whether the effect was due to on- or off-target effects of the compound. In order to determine the mechanism of action, it may be necessary to quantify the specificity of the inhibitor, a nontrivial task. To address this challenge, Bantscheff *et al.* [58] developed a kinase capturing bead ("Kinobead") consisting of multiple immobilized broad-selectivity kinase inhibitors. On application to cell lysate, a large number of kinases (and other purine-binding proteins) are retained on the Kinobeads due to the interaction with the kinase inhibitors. To obtain a quantitative target profile of a selected compound, cells or cell lysate are treated with the compound at varying concentrations prior to affinity isolation with the Kinobead. Kinases inhibited by the selected compound exhibit decreased binding to the Kinobead and therefore yield decreased signal by quantitative (iTRAQ-based) MS. Combining this approach with phosphorylation analysis can yield a profile as to the phosphorylation status of the kinases bound to the selected compound, potentially identifying whether the compound binds to the active or inactive isoform of the kinase. After establishing the specificity of the inhibitor, it will then be possible to regather quantitative phosphoproteomic data to determine the effect on the cell signaling network of inhibiting the selected targets of the inhibitor. The workflow for this approach is outlined schematically in Fig. 3. Following through this iterative process, one could begin to build out downstream signaling networks directly or indirectly affected by a selected kinase in the context of various human pathologies.

11 Conclusions

What does the future hold for quantitative phosphoproteomics by MS? The field is in a rapid state of flux, including new enrichment strategies, novel quantification reagents, and new instrumentation. With each improvement it becomes possible to identify and quantify increasing numbers of phosphorylation sites, digging deeper and deeper into the elusive comprehensive phosphoproteome. However, as many of the above applications demonstrate, size of the dataset is not always the most important metric. Instead, understanding the biological implications of many of the phosphorylation sites is critical, since the ultimate goal of most of these studies is to increase insight into cellular signaling and biological control. Linking phosphorylation data to other quantitative phenotypic endpoints is a crucial step in this procedure, and one that has been often ignored in the effort to gather larger data sets. Going forward, the combi-

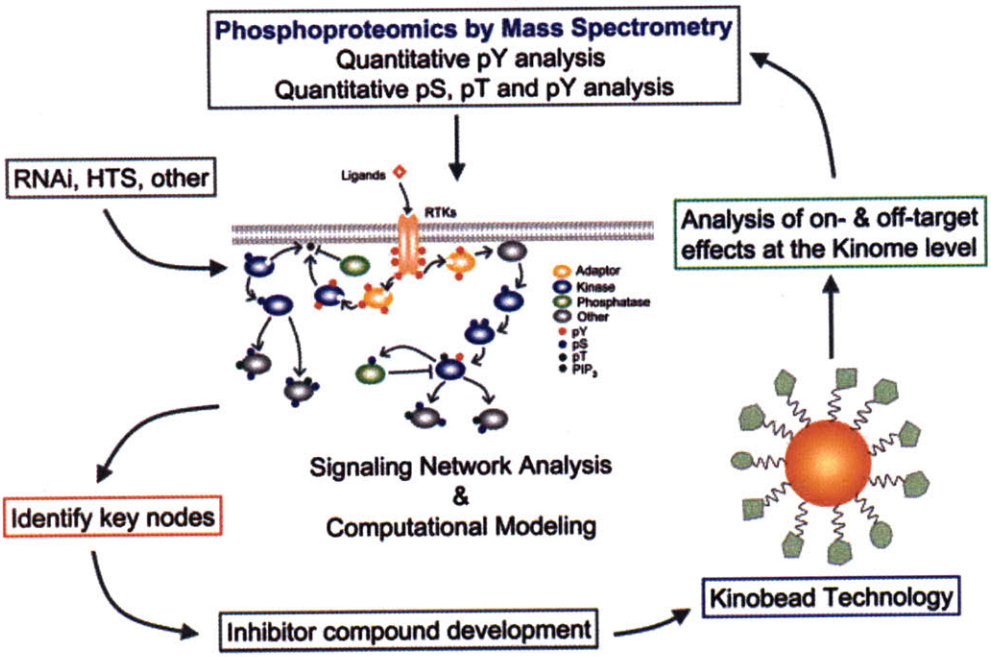


Figure 3. Signaling network analysis by mass spectrometry, drug-target discovery, and kinase inhibition. Mass spectrometry-based phosphoproteomics is used to quantify signaling networks in various cell states resulting from multiple perturbations (*e.g.*, cytokine stimulation, RNAi, small molecule inhibitors). Computational modeling is used to link this information to quantitative phenotypic measurements, thereby identifying key nodes in the signaling network. To validate this model, compounds are developed to target these nodes, and the phenotypic effect is monitored. The specificity of the compounds can then be tested by the Kinobead assay to quantify on- and off-target effects. Changes in the signaling network can then be quantified to establish kinase–substrate relationships given the established targets of the compound.

nation of MS, phenotypic characterization, mathematical modeling, and selected perturbations should provide rapid advancement in our understanding of the complexities of cellular signaling network, information that will enable the development of better therapeutic agents with fewer off-target effects.

The authors have declared no conflict of interest.

12 References

[1] Hunter, T., Signaling–2000 and beyond. *Cell* 2000, *100*, 113–127.

[2] Seet, B. T., Dikic, I., Zhou, M. M., Pawson, T., Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* 2006, *7*, 473–483.

[3] Huang, P. H., Cavenee, W. K., Furnari, F. B., White, F. M., Uncovering therapeutic targets for glioblastoma: A systems biology approach. *Cell Cycle* 2007, *6*, 2750–2754.

[4] Pawson, T., Linding, R., Network medicine. *FEBS Lett.* 2008.

[5] Pandey, A., Podtelejnikov, A. V., Blagoev, B., Bustelo, X. R. *et al.*, Analysis of receptor signaling pathways by mass spectrometry: Identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 179–184.

[6] Andersson, L., Porath, J., Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal. Biochem.* 1986, *154*, 250–254.

[7] Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E. *et al.*, Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 12130–12135.

[8] Rush, J., Moritz, A., Lee, K. A., Guo, A. *et al.*, Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* 2004.

[9] Blagoev, B., Ong, S. E., Kratchmarova, I., Mann, M., Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* 2004, *22*, 1139–1145.

[10] Hinsby, A. M., Olsen, J. V., Bennett, K. L., Mann, M., Signaling initiated by overexpression of the fibroblast growth factor receptor-1 investigated by mass spectrometry. *Mol. Cell. Proteomics* 2003, *2*, 29–36.

[11] Ibarrola, N., Molina, H., Iwahori, A., Pandey, A., A novel proteomic approach for specific identification of tyrosine kinase substrates using 13C-labeled tyrosine. *J. Biol. Chem.* 2004.

[12] Hinsby, A. M., Olsen, J. V., Mann, M., Tyrosine phosphoproteomics of fibroblast growth factor signaling: A role for insulin receptor substrate-4. *J. Biol. Chem.* 2004, *279*, 46438–46447.

[13] Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., III *et al.*, ATM and ATR substrate analysis reveals exten-

sive protein networks responsive to DNA damage. *Science* 2007, *316*, 1160–1166.

[14] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. *et al.*, Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 2002, *20*, 301–305.

[15] Ficarro, S. B., Salomon, A. R., Brill, L. M., Mason, D. E. *et al.*, Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun. Mass Spectrom.* 2005, *19*, 57–71.

[16] Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., Jorgensen, T. J., Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 2005, *4*, 873–886.

[17] Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P. *et al.*, Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* 2005, *4*, 310–327.

[18] Ballif, B. A., Villen, J., Beausoleil, S. A., Schwartz, D., Gygi, S. P., Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* 2004, *3*, 1093–1101.

[19] Motoyama, A., Xu, T., Ruse, C. I., Wohlschlegel, J. A., Yates, J. R., III, Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides. *Anal. Chem.* 2007, *79*, 3623–3634.

[20] Oda, Y., Nagasu, T., Chait, B. T., Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* 2001, *19*, 379–382.

[21] Tao, W. A., Wollscheid, B., O'Brien, R., Eng, J. K. *et al.*, Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. *Nat. Methods* 2005, *2*, 591–598.

[22] Bodenmiller, B., Mueller, L. N., Pedrioli, P. G., Pflieger, D. *et al.*, An integrated chemical, mass spectrometric and computational strategy for (quantitative) phosphoproteomics: Application to *Drosophila melanogaster* Kc167 cells. *Mol. Biosyst.* 2007, *3*, 275–286.

[23] Schmelzle, K., Kane, S., Gridley, S., Lienhard, G. E., White, F. M., Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* 2006, *55*, 2171–2179.

[24] Kim, J. E., White, F. M., Quantitative analysis of phosphotyrosine signaling networks triggered by CD3 and CD28 costimulation in Jurkat cells. *J. Immunol.* 2006, *176*, 2833–2843.

[25] Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S. *et al.*, Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Mol. Syst. Biol.* 2006, *2*, 54.

[26] Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J. *et al.*, Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* 2005, *4*, 1240–1250.

[27] Choe, L., D'Ascenzo, M., Relkin, N. R., Pappin, D. *et al.*, 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* 2007, *7*, 3651–3660.

[28] Ong, S. E., Mann, M., A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* 2006, *1*, 2650–2660.

[29] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B. *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 2002, *1*, 376–386.

[30] Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T. *et al.*, Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 2005, *122*, 209–220.

[31] Nirmalan, N., Sims, P. F., Hyde, J. E., Quantitative proteomics of the human malaria parasite *Plasmodium falciparum* and its application to studies of development and inhibition. *Mol. Microbiol.* 2004, *52*, 1187–1199.

[32] Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M., Jensen, O. N., Stable isotope labeling of *Arabidopsis thaliana* cells and quantitative proteomics by mass spectrometry. *Mol. Cell. Proteomics* 2005, *4*, 1697–1709.

[33] Krijgsveld, J., Ketting, R. F., Mahmoudi, T., Johansen, J. *et al.*, Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nat. Biotechnol.* 2003, *21*, 927–931.

[34] Wu, C. C., MacCoss, M. J., Howell, K. E., Matthews, D. E., Yates, J. R., III, Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* 2004, *76*, 4951–4959.

[35] Barnidge, D. R., Hall, G. D., Stocker, J. L., Muddiman, D. C., Evaluation of a cleavable stable isotope labeled synthetic peptide for absolute protein quantification using LC-MS/MS. *J. Proteome Res.* 2004, *3*, 658–661.

[36] Beynon, R. J., Doherty, M. K., Pratt, J. M., Gaskell, S. J., Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* 2005, *2*, 587–589.

[37] Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J. *et al.*, Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat. Protoc.* 2006, *1*, 1029–1043.

[38] Mirzaei, H., McBee, J., Watts, J., Aebersold, R., Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. *Mol. Cell. Proteomics* 2007, *7*, 813–823.

[39] Foster, L. J., de Hoog, C. L., Zhang, Y., Zhang, Y. *et al.*, A mammalian organelle map by protein correlation profiling. *Cell* 2006, *125*, 187–199.

[40] Ishihama, Y., Oda, Y., Tabata, T., Sato, T. *et al.*, Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* 2005, *4*, 1265–1272.

[41] Mallick, P., Schirle, M., Chen, S. S., Flory, M. R. *et al.*, Computational prediction of proteotypic peptides for quantitative proteomics. *Nat. Biotechnol.* 2007, *25*, 125–131.

[42] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F. *et al.*, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 1999, *17*, 994–999.

[43] Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L. *et al.*, Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* 2000, *72*, 563–573.

[44] Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 9528–9533.

[45] Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A., White, F. M., Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 5860–5865.

[46] Olsen, J. V., Macek, B., Lange, O., Makarov, A. *et al.*, Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* 2007, *4*, 709–712.

[47] Yarden, Y., Sliwkowski, M. X., Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* 2001, *2*, 127–137.

[48] Rikova, K., Guo, A., Zeng, Q., Possemato, A. *et al.*, Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 2007, *131*, 1190–1203.

[49] Soda, M., Choi, Y. L., Enomoto, M., Takada, S. *et al.*, Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007, *448*, 561–566.

[50] Olsen, J. V., Blagoev, B., Gnad, F., Macek, B. *et al.*, Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006, *127*, 635–648.

[51] Yaffe, M. B., White, F. M., Signaling netwErks get the global treatment. *Genome Biol.* 2007, *8*, 202.

[52] Friedman, A., Perrimon, N., A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling. *Nature* 2006, *444*, 230–234.

[53] Elias, J. E., Haas, W., Faherty, B. K., Gygi, S. P., Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat. Methods* 2005, *2*, 667–675.

[54] Kumar, N., Wolf-Yadlin, A., White, F. M., Lauffenburger, D. A., Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput. Biol.* 2007, *3*, e4.

[55] Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., Mann, M., Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* 2005, *308*, 1472–1477.

[56] Huang, P. H., Mukasa, A., Bonavia, R., Flynn, R. A. *et al.*, Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 12867–12872.

[57] Stommel, J. M., Kimmelman, A. C., Ying, H., Nabioullin, R. *et al.*, Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. *Science* 2007, *318*, 287–290.

[58] Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S. *et al.*, Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat. Biotechnol.* 2007, *25*, 1035–1044.

[59] Zhou, H., Watts, J. D., Aebersold, R., A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* 2001, *19*, 375–378.

[60] Chi, A., Huttenhower, C., Geer, L. Y., Coon, J. J. *et al.*, Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 2193–2198.

I.4 Motivations

Tamoxifen treatment has been beneficial to millions of hormone receptor positive breast cancer patients in the past several decades. The therapeutic regimens are effective in prolonging the disease free survival period, cost-efficient, and maintain high level of quality of life (QOL) for the majority of hormone receptor positive breast cancer patients. It is regarded as the one of the best cancer therapeutics available world-wide for women's health and welfare. However, even such a successful therapeutic compound is not perfect. Due to the large number of breast cancer patients who are put on the Tamoxifen treatment, any small problem regarding its use could affect tremendous number of breast cancer patients worldwide. In this regard, Tamoxifen resistance is a significant problem that deteriorates breast cancer patients' QOL due to aggressive tumors that require severe chemotherapies, giving breast cancer patients a higher degree of burden and suffering.

In order to address the Tamoxifen resistance problem, we need to better understand the disease mechanisms. Only through a better understanding of the

disease mechanisms is the better development of new therapeutics feasible. Our laboratory has previously developed a quantitative chemical proteomic approach to dissect phosphorylation signaling events in a system-wide fashion and has successfully applied it to the study of the cancer model systems such as EGFR vIII signaling networks in human glioblastoma cell lines. In this thesis, I have applied the phosphoproteomic analysis by mass spectrometry, which is the core of Forest White Lab's technology, to the study of breast cancer cell line models with the Tamoxifen resistance.

In order to study Tamoxifen resistance in breast cancer cell line models with certain degrees of the clinical relevance, we have employed the classical breast cancer cell line used for development of many current endocrine therapies. Such models cell lines are modified to obtain two Tamoxifen resistant cell line models. The two Tamoxifen resistant model cell lines are previously documented by multiple groups and have been valuable cell line models to investigate the disease mechanisms of the Tamoxifen resistance.

Although much work has been done in the past to elucidate genetic changes that cause Tamoxifen resistance, the disease mechanisms that breast

cancer cells employ to become anti-estrogen resistant are still not completely understood. Furthermore, there are several important questions in which there is a lack of consensus in the field, including if HER2 overexpression confers Tamoxifen resistance in all $ER^+/HER2^+$ breast cancer patients, issues of ER's non-genomic functions, and the cross-talk in $ER^+/HER2^+$ systems.

The aim of this doctoral thesis is mainly three-fold. Firstly, the aim is to employ quantitative phosphoproteomic analysis based on mass spectrometry to gain a better understanding of the signaling changes due to exposure to Tamoxifen treatment. Secondly, the aim is to compare the differences in signaling networks between the Tamoxifen sensitive cells and the Tamoxifen resistant cells in order to identify promising therapeutic targets. Thirdly, the aim is to test the hypotheses generated from the quantitative phosphoproteomic analysis in model cell line systems in order to validate the importance of proposed therapeutic targets.

This thesis incorporates the experimental results including development of the Tamoxifen resistant cell lines from the Tamoxifen sensitive cell line, obtaining quantitative phosphotyrosine data, the treatment of model cell lines with small molecule inhibitors, phenotypic analyses such as migration assays and invasion

assays, rescue experiments with growth factor treatments, and finally the mass-spec analysis of human breast tumor specimens. The analysis has revealed interesting insights into network modulation and how the Tamoxifen resistant cells fine-tune the phosphorylation levels of pro-survival and proliferative signals. In the course of this thesis, I have established the power of quantitative phosphoproteomics as a tool to dissect the therapeutic-resistance mechanisms by analysing the phosphotyrosine mediated signaling network and as a means for identifying molecular targets for pharmacological therapeutic interventions.

References

Anderson, W.F., Chatterjee, N., Ershler, W.B., and Brawley, O.W. (2002). Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database. Breast cancer research and treatment *76*, 27-36.

Argiris, A., Wang, C.X., Whalen, S.G., and DiGiovanna, M.P. (2004). Synergistic interactions between tamoxifen and trastuzumab (Herceptin). Clin Cancer Res *10*, 1409-1420.

Ballif, B.A., Villen, J., Beausoleil, S.A., Schwartz, D., and Gygi, S.P. (2004). Phosphoproteomic analysis of the developing mouse brain. Mol Cell Proteomics *3*, 1093-1101.

Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., *et al.* (2007). Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. Nat Biotechnol *25*, 1035-1044.

Barnidge, D.R., Hall, G.D., Stocker, J.L., and Muddiman, D.C. (2004). Evaluation of a cleavable stable isotope labeled synthetic peptide for absolute protein quantification using LC-MS/MS. J Proteome Res *3*, 658-661.

Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C., and Gygi, S.P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. Proceedings of the National Academy of Sciences of the United States of America *101*, 12130-12135.

Benz, C.C., Scott, G.K., Sarup, J.C., Johnson, R.M., Tripathy, D., Coronado, E., Shepard, H.M., and Osborne, C.K. (1992). Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. Breast cancer research and treatment *24*, 85-95.

Beynon, R.J., Doherty, M.K., Pratt, J.M., and Gaskell, S.J. (2005). Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods *2*, 587-589.

Bjornstrom, L., and Sjoberg, M. (2005). Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. Molecular endocrinology (Baltimore, Md *19*, 833-842.

Blagoev, B., Ong, S.E., Kratchmarova, I., and Mann, M. (2004). Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol *22*, 1139-1145.

Bodenmiller, B., Mueller, L.N., Pedrioli, P.G., Pflieger, D., Junger, M.A., Eng, J.K., Aebersold, R., and Tao, W.A. (2007). An integrated chemical, mass spectrometric and computational strategy for (quantitative) phosphoproteomics: application to Drosophila melanogaster Kc167 cells. Molecular bioSystems *3*, 275-286.

Bonneterre, J., Thurlimann, B., Robertson, J.F., Krzakowski, M., Mauriac, L., Koralewski, P., Vergote, I., Webster, A., Steinberg, M., and von Euler, M. (2000). Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the Tamoxifen or Arimidex Randomized Group Efficacy and Tolerability study. J Clin Oncol *18*, 3748-3757.

Boulay, A., Rudloff, J., Ye, J., Zumstein-Mecker, S., O'Reilly, T., Evans, D.B., Chen, S., and Lane, H.A. (2005). Dual inhibition of mTOR and estrogen receptor signaling in vitro induces cell death in models of breast cancer. Clin Cancer Res *11*, 5319-5328.

Buchholz, T.A., Hunt, K.K., Amosson, C.M., Tucker, S.L., Strom, E.A., McNeese, M.D., Buzdar, A.U., Singletary, S.E., and Hortobagyi, G.N. (1999). Sequencing of chemotherapy and radiation in lymph node-negative breast cancer. Cancer J Sci Am *5*, 159-164.

Burnett, G., and Kennedy, E.P. (1954). The enzymatic phosphorylation of proteins. The Journal of biological chemistry *211*, 969-980.

Campbell, R.A., Bhat-Nakshatri, P., Patel, N.M., Constantinidou, D., Ali, S., and Nakshatri, H. (2001). Phosphatidylinositol 3-kinase/AKT-mediated activation of estrogen receptor alpha: a new model for anti-estrogen resistance. The Journal of biological chemistry *276*, 9817-9824.

cancer, N.C.C.N.B. (Accessed March 1st 2009). Clinical practice guidelines in oncology. *version I 2009*.

Carlomagno, C., Perrone, F., Gallo, C., De Laurentiis, M., Lauria, R., Morabito, A., Pettinato, G., Panico, L., D'Antonio, A., Bianco, A.R., *et al.* (1996). c-erb B2 overexpression decreases the benefit of adjuvant tamoxifen in early-stage breast cancer without axillary lymph node metastases. J Clin Oncol *14*, 2702-2708.

Chan, C.M., Martin, L.A., Johnston, S.R., Ali, S., and Dowsett, M. (2002). Molecular changes associated with the acquisition of oestrogen hypersensitivity in MCF-7 breast cancer cells on long-term oestrogen deprivation. The Journal of steroid biochemistry and molecular biology *81*, 333-341.

Choe, L., D'Ascenzo, M., Relkin, N.R., Pappin, D., Ross, P., Williamson, B., Guertin, S., Pribil, P., and Lee, K.H. (2007). 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. Proteomics *7*, 3651-3660.

Chung, Y.L., Sheu, M.L., Yang, S.C., Lin, C.H., and Yen, S.H. (2002). Resistance to tamoxifen-induced apoptosis is associated with direct interaction between Her2/neu and cell membrane estrogen receptor in breast cancer. International journal of cancer *97*, 306-312.

Coates, A.S., Keshaviah, A., Thurlimann, B., Mouridsen, H., Mauriac, L., Forbes, J.F., Paridaens, R., Castiglione-Gertsch, M., Gelber, R.D., Colleoni, M., *et al.* (2007).

Five years of letrozole compared with tamoxifen as initial adjuvant therapy for postmenopausal women with endocrine-responsive early breast cancer: update of study BIG 1-98. J Clin Oncol *25*, 486-492.

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation--a 25 year update. Trends in biochemical sciences *25*, 596-601.

Coombes, R.C., Kilburn, L.S., Snowdon, C.F., Paridaens, R., Coleman, R.E., Jones, S.E., Jassem, J., Van de Velde, C.J., Delozier, T., Alvarez, I., *et al.* (2007). Survival and safety of exemestane versus tamoxifen after 2-3 years' tamoxifen treatment (Intergroup Exemestane Study): a randomised controlled trial. Lancet *369*, 559-570.

De Placido, S., De Laurentiis, M., Carlomagno, C., Gallo, C., Perrone, F., Pepe, S., Ruggiero, A., Marinelli, A., Pagliarulo, C., Panico, L., *et al.* (2003). Twenty-year results of the Naples GUN randomized trial: predictive factors of adjuvant tamoxifen efficacy in early breast cancer. Clin Cancer Res *9*, 1039-1046.

Dowsett, M., Allred, C., Knox, J., Quinn, E., Salter, J., Wale, C., Cuzick, J., Houghton, J., Williams, N., Mallon, E., *et al.* (2008). Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the Arimidex, Tamoxifen, Alone or in Combination trial. J Clin Oncol *26*, 1059-1065.

Eiermann, W., Paepke, S., Appfelstaedt, J., Llombart-Cussac, A., Eremin, J., Vinholes, J., Mauriac, L., Ellis, M., Lassus, M., Chaudri-Ross, H.A., *et al.* (2001). Preoperative treatment of postmenopausal breast cancer patients with letrozole: A randomized double-blind multicenter study. Ann Oncol *12*, 1527-1532.

Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat Methods *2*, 667-675.

Elledge, R.M., Green, S., Ciocca, D., Pugh, R., Allred, D.C., Clark, G.M., Hill, J.,

Ravdin, P., O'Sullivan, J., Martino, S., *et al.* (1998). HER-2 expression and response to tamoxifen in estrogen receptor-positive breast cancer: a Southwest Oncology Group Study. Clin Cancer Res *4*, 7-12.

Ellis, C.A., Vos, M.D., Wickline, M., Riley, C., Vallecorsa, T., Telford, W.G., Zujewskil, J., and Clark, G.J. (2003). Tamoxifen and the farnesyl transferase inhibitor FTI-277 synergize to inhibit growth in estrogen receptor-positive breast tumor cell lines. Breast cancer research and treatment *78*, 59-67.

Ellis, M.J., Coop, A., Singh, B., Mauriac, L., Llombert-Cussac, A., Janicke, F., Miller, W.R., Evans, D.B., Dugan, M., Brady, C., *et al.* (2001). Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. J Clin Oncol *19*, 3808-3816.

Ellis, M.J., and Ma, C. (2007). Letrozole in the neoadjuvant setting: the P024 trial. Breast cancer research and treatment *105 Suppl 1*, 33-43.

Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nature biotechnology *20*, 301-305.

Fisher, B., Costantino, J.P., Wickerham, D.L., Redmond, C.K., Kavanah, M., Cronin, W.M., Vogel, V., Robidoux, A., Dimitrov, N., Atkins, J., *et al.* (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst *90*, 1371-1388.

Fisher, B., Dignam, J., Bryant, J., DeCillis, A., Wickerham, D.L., Wolmark, N., Costantino, J., Redmond, C., Fisher, E.R., Bowman, D.M., *et al.* (1996). Five versus more than five years of tamoxifen therapy for breast cancer patients with negative lymph nodes and estrogen receptor-positive tumors. J Natl Cancer Inst *88*, 1529-1542.

Forbes, J.F., Cuzick, J., Buzdar, A., Howell, A., Tobias, J.S., and Baum, M. (2008). Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. The lancet oncology *9*, 45-53.

Foster, L.J., de Hoog, C.L., Zhang, Y., Zhang, Y., Xie, X., Mootha, V.K., and Mann, M. (2006). A mammalian organelle map by protein correlation profiling. Cell *125*, 187-199.

Fradelizi, J., Friederich, E., Beckerle, M.C., and Golsteyn, R.M. (1999). Quantitative measurement of proteins by western blotting with Cy5-coupled secondary antibodies. BioTechniques *26*, 484-486, 488, 490 passim.

Friedman, A., and Perrimon, N. (2006). A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling. Nature *444*, 230-234.

Gee, J.M., Harper, M.E., Hutcheson, I.R., Madden, T.A., Barrow, D., Knowlden, J.M., McClelland, R.A., Jordan, N., Wakeling, A.E., and Nicholson, R.I. (2003). The antiepidermal growth factor receptor agent gefitinib (ZD1839/Iressa) improves antihormone response and prevents development of resistance in breast cancer in vitro. Endocrinology *144*, 5105-5117.

Gembitsky, D.S., Lawlor, K., Jacovina, A., Yaneva, M., and Tempst, P. (2004). A prototype antibody microarray platform to monitor changes in protein tyrosine phosphorylation. Mol Cell Proteomics *3*, 1102-1118.

Gingrich, J.C., Davis, D.R., and Nguyen, Q. (2000). Multiplex detection and quantitation of proteins on western blots using fluorescent probes. BioTechniques *29*, 636-642.

Gruhler, A., Olsen, J.V., Mohammed, S., Mortensen, P., Faergeman, N.J., Mann, M., and Jensen, O.N. (2005a). Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol Cell Proteomics *4*, 310-327.

Gruhler, A., Schulze, W.X., Matthiesen, R., Mann, M., and Jensen, O.N. (2005b). Stable isotope labeling of Arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. Mol Cell Proteomics *4*, 1697-1709.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol *17*, 994-999.

Hinsby, A.M., Olsen, J.V., Bennett, K.L., and Mann, M. (2003). Signaling initiated by overexpression of the fibroblast growth factor receptor-1 investigated by mass spectrometry. Mol Cell Proteomics *2*, 29-36.

Hinsby, A.M., Olsen, J.V., and Mann, M. (2004). Tyrosine phosphoproteomics of fibroblast growth factor signaling: a role for insulin receptor substrate-4. The Journal of biological chemistry *279*, 46438-46447.

Huang, P.H., Mukasa, A., Bonavia, R., Flynn, R.A., Brewer, Z.E., Cavenee, W.K., Furnari, F.B., and White, F.M. (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. Proc Natl Acad Sci U S A *104*, 12867-12872.

Hunter, T. (2000). Signaling--2000 and beyond. Cell *100*, 113-127.

Hunter, T., and Sefton, B.M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proceedings of the National Academy of Sciences of the United States of America *77*, 1311-1315.

Ibarrola, N., Molina, H., Iwahori, A., and Pandey, A. (2004). A novel proteomic approach for specific identification of tyrosine kinase substrates using [13C]tyrosine. The Journal of biological chemistry *279*, 15805-15813.

Irish, J.M., Hovland, R., Krutzik, P.O., Perez, O.D., Bruserud, O., Gjertsen, B.T., and Nolan, G.P. (2004). Single cell profiling of potentiated phospho-protein networks in

cancer cells. Cell *118*, 217-228.

Irish, J.M., Kotecha, N., and Nolan, G.P. (2006). Mapping normal and cancer cell signalling networks: towards single-cell proteomics. Nature reviews *6*, 146-155.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics *4*, 1265-1272.

Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., Smigal, C., and Thun, M.J. (2006). Cancer Statistics, 2006, C.J. Clin, ed., pp. 106-130.

Jensen, E.V., Greene, G.L., Closs, L.E., DeSombre, E.R., and Nadji, M. (1982). Receptors reconsidered: a 20-year perspective. Recent Prog Horm Res *38*, 1-40.

Joel, P.B., Smith, J., Sturgill, T.W., Fisher, T.L., Blenis, J., and Lannigan, D.A. (1998). pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. Molecular and cellular biology *18*, 1978-1984.

Johnston, S.R. (1997). Acquired tamoxifen resistance in human breast cancer--potential mechanisms and clinical implications. Anti-cancer drugs *8*, 911-930.

Johnston, S.R. (2005). Combinations of endocrine and biological agents: present status of therapeutic and presurgical investigations. Clin Cancer Res *11*, 889s-899s.

Joughin, B.A., Naegle, K.M., Huang, P.H., Yaffe, M.B., Lauffenburger, D.A., and White, F.M. (2009). An integrated comparative phosphoproteomic and bioinformatic approach reveals a novel class of MPM-2 motifs upregulated in EGFRvIII-expressing glioblastoma cells. Molecular bioSystems *5*, 59-67.

Kahlert, S., Nuedling, S., van Eickels, M., Vetter, H., Meyer, R., and Grohe, C.

(2000). Estrogen receptor alpha rapidly activates the IGF-1 receptor pathway. The Journal of biological chemistry *275*, 18447-18453.

Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., Masushige, S., Gotoh, Y., Nishida, E., Kawashima, H., *et al.* (1995). Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. Science (New York, NY *270*, 1491-1494.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., *et al.* (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209-220.

Kim, J.E., and White, F.M. (2006). Quantitative analysis of phosphotyrosine signaling networks triggered by CD3 and CD28 costimulation in Jurkat cells. J Immunol *176*, 2833-2843.

Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., and Mann, M. (2005). Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. Science *308*, 1472-1477.

Krijgsveld, J., Ketting, R.F., Mahmoudi, T., Johansen, J., Artal-Sanz, M., Verrijzer, C.P., Plasterk, R.H., and Heck, A.J. (2003). Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. Nat Biotechnol *21*, 927-931.

Kumar, N., Wolf-Yadlin, A., White, F.M., and Lauffenburger, D.A. (2007). Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. PLoS Comput Biol *3*, e4.

Kushner, P.J., Agard, D.A., Greene, G.L., Scanlan, T.S., Shiau, A.K., Uht, R.M., and Webb, P. (2000). Estrogen receptor pathways to AP-1. The Journal of steroid biochemistry and molecular biology *74*, 311-317.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics *4*, 873-886.

Levin, E.R. (2003). Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor. Molecular endocrinology (Baltimore, Md *17*, 309-317.

Lewis, J.S., and Jordan, V.C. (2005). Selective estrogen receptor modulators (SERMs): mechanisms of anticarcinogenesis and drug resistance. Mutat Res *591*, 247-263.

Lipton, A., Ali, S.M., Leitzel, K., Demers, L., Harvey, H.A., Chaudri-Ross, H.A., Brady, C., Wyld, P., and Carney, W. (2003). Serum HER-2/neu and response to the aromatase inhibitor letrozole versus tamoxifen. J Clin Oncol *21*, 1967-1972.

Liu, G., Marrinan, C.H., Taylor, S.A., Black, S., Basso, A.D., Kirschmeier, P., Robert Bishop, W., Liu, M., and Long, B.J. (2007). Enhancement of the antitumor activity of tamoxifen and anastrozole by the farnesyltransferase inhibitor lonafarnib (SCH66336). Anti-cancer drugs *18*, 923-931.

MacBeath, G. (2002). Protein microarrays and proteomics. Nature genetics *32 Suppl*, 526-532.

Mallick, P., Schirle, M., Chen, S.S., Flory, M.R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., *et al.* (2007). Computational prediction of proteotypic peptides for quantitative proteomics. Nat Biotechnol *25*, 125-131.

Marcom, P.K., Isaacs, C., Harris, L., Wong, Z.W., Kommarreddy, A., Novielli, N., Mann, G., Tao, Y., and Ellis, M.J. (2007). The combination of letrozole and trastuzumab as first or second-line biological therapy produces durable responses in a subset of HER2 positive and ER positive advanced breast cancers. Breast cancer research and treatment *102*, 43-49.

Masamura, S., Santner, S.J., Heitjan, D.F., and Santen, R.J. (1995). Estrogen deprivation causes estradiol hypersensitivity in human breast cancer cells. J Clin Endocrinol Metab *80*, 2918-2925.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., *et al.* (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science (New York, NY *316*, 1160-1166.

Migliaccio, A., Di Domenico, M., Castoria, G., de Falco, A., Bontempo, P., Nola, E., and Auricchio, F. (1996). Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. The EMBO journal *15*, 1292-1300.

Mirzaei, H., McBee, J., Watts, J., and Aebersold, R. (2007). Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. Mol Cell Proteomics.

Morseman, J.P., Moss, M.W., Zoha, S.J., and Allnutt, F.C. (1999). PBXL-1: a new fluorochrome applied to detection of proteins on membranes. BioTechniques *26*, 559-563.

Moser, K., and White, F.M. (2006). Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. Journal of proteome research *5*, 98-104.

Motoyama, A., Xu, T., Ruse, C.I., Wohlschlegel, J.A., and Yates, J.R., 3rd (2007). Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides. Analytical chemistry *79*, 3623-3634.

Mouridsen, H., Gershanovich, M., Sun, Y., Perez-Carrion, R., Boni, C., Monnier, A., Apffelstaedt, J., Smith, R., Sleeboom, H.P., Jaenicke, F., *et al.* (2003). Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group. J Clin Oncol *21*, 2101-2109.

Nabholtz, J.M., Buzdar, A., Pollak, M., Harwin, W., Burton, G., Mangalik, A., Steinberg, M., Webster, A., and von Euler, M. (2000). Anastrozole is superior to tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: results of a North American multicenter randomized trial. Arimidex Study Group. J Clin Oncol *18*, 3758-3767.

Nielsen, U.B., Cardone, M.H., Sinskey, A.J., MacBeath, G., and Sorger, P.K. (2003). Profiling receptor tyrosine kinase activation by using Ab microarrays. Proceedings of the National Academy of Sciences of the United States of America *100*, 9330-9335.

Nirmalan, N., Sims, P.F., and Hyde, J.E. (2004). Quantitative proteomics of the human malaria parasite Plasmodium falciparum and its application to studies of development and inhibition. Mol Microbiol *52*, 1187-1199.

Normanno, N., Di Maio, M., De Maio, E., De Luca, A., de Matteis, A., Giordano, A., and Perrone, F. (2005). Mechanisms of endocrine resistance and novel therapeutic strategies in breast cancer. Endocr Relat Cancer *12*, 721-747.

Oda, Y., Nagasu, T., and Chait, B.T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nature biotechnology *19*, 379-382.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635-648.

Olsen, J.V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods *4*, 709-712.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC,

as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Ong, S.E., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc *1*, 2650-2660.

Osborne, C.K., and Schiff, R. (2005). Estrogen-receptor biology: continuing progress and therapeutic implications. J Clin Oncol *23*, 1616-1622.

Paridaens, R., Dirix, L., Lohrisch, C., Beex, L., Nooij, M., Cameron, D., Biganzoli, L., Cufer, T., Duchateau, L., Hamilton, A., *et al.* (2003). Mature results of a randomized phase II multicenter study of exemestane versus tamoxifen as first-line hormone therapy for postmenopausal women with metastatic breast cancer. Ann Oncol *14*, 1391-1398.

Perez, O.D., and Nolan, G.P. (2006). Phospho-proteomic immune analysis by flow cytometry: from mechanism to translational medicine at the single-cell level. Immunological reviews *210*, 208-228.

Polychronis, A., Sinnett, H.D., Hadjiminas, D., Singhal, H., Mansi, J.L., Shivapatham, D., Shousha, S., Jiang, J., Peston, D., Barrett, N., *et al.* (2005). Preoperative gefitinib versus gefitinib and anastrozole in postmenopausal patients with oestrogen-receptor positive and epidermal-growth-factor-receptor-positive primary breast cancer: a double-blind placebo-controlled phase II randomised trial. The lancet oncology *6*, 383-391.

Pratt, J.M., Simpson, D.M., Doherty, M.K., Rivers, J., Gaskell, S.J., and Beynon, R.J. (2006). Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. Nat Protoc *1*, 1029-1043.

Razandi, M., Pedram, A., Park, S.T., and Levin, E.R. (2003). Proximal events in signaling by plasma membrane estrogen receptors. The Journal of biological chemistry *278*, 2701-2712.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., *et al.* (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell *131*, 1190-1203.

Ropero, S., Menendez, J.A., Vazquez-Martin, A., Montero, S., Cortes-Funes, H., and Colomer, R. (2004). Trastuzumab plus tamoxifen: anti-proliferative and molecular interactions in breast carcinoma. Breast cancer research and treatment *86*, 125-137.

Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nature biotechnology *23*, 94-101.

Santen, R.J., Song, R.X., Zhang, Z., Kumar, R., Jeng, M.H., Masamura, S., Lawrence, J., Jr., MacMahon, L.P., Yue, W., and Berstein, L. (2005). Adaptive hypersensitivity to estrogen: mechanisms and clinical relevance to aromatase inhibitor therapy in breast cancer treatment. The Journal of steroid biochemistry and molecular biology *95*, 155-165.

Santen, R.J., Song, R.X., Zhang, Z., Kumar, R., Jeng, M.H., Masamura, S., Yue, W., and Berstein, L. (2003). Adaptive hypersensitivity to estrogen: mechanism for superiority of aromatase inhibitors over selective estrogen receptor modulators for breast cancer treatment and prevention. Endocr Relat Cancer *10*, 111-130.

Schiff, R., Massarweh, S.A., Shou, J., Bharwani, L., Mohsin, S.K., and Osborne, C.K. (2004). Cross-talk between estrogen receptor and growth factor pathways as a molecular target for overcoming endocrine resistance. Clin Cancer Res *10*, 331S-336S.

Schmelzle, K., Kane, S., Gridley, S., Lienhard, G.E., and White, F.M. (2006). Temporal dynamics of tyrosine phosphorylation in insulin signaling. Diabetes *55*, 2171-2179.

Shou, J., Massarweh, S., Osborne, C.K., Wakeling, A.E., Ali, S., Weiss, H., and Schiff, R. (2004). Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. J Natl Cancer Inst *96*, 926-935.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A., and McGuire, W.L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science (New York, NY *235*, 177-182.

Smith, I.E., and Dowsett, M. (2003). Aromatase inhibitors in breast cancer. N Engl J Med *348*, 2431-2442.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., *et al.* (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561-566.

Stommel, J.M., Kimmelman, A.C., Ying, H., Nabioullin, R., Ponugoti, A.H., Wiedemeyer, R., Stegh, A.H., Bradner, J.E., Ligon, K.L., Brennan, C., *et al.* (2007). Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. Science *318*, 287-290.

Sun, M., Paciga, J.E., Feldman, R.I., Yuan, Z., Coppola, D., Lu, Y.Y., Shelley, S.A., Nicosia, S.V., and Cheng, J.Q. (2001). Phosphatidylinositol-3-OH Kinase (PI3K)/AKT2, activated in breast cancer, regulates and is induced by estrogen receptor alpha (ERalpha) via interaction between ERalpha and PI3K. Cancer research *61*, 5985-5991.

Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., and Hunt, D.F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A *101*, 9528-9533.

Tao, W.A., Wollscheid, B., O'Brien, R., Eng, J.K., Li, X.J., Bodenmiller, B., Watts,

J.D., Hood, L., and Aebersold, R. (2005). Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. Nature methods *2*, 591-598.

White, F.M. (2008). Quantitative phosphoproteomic analysis of signaling network dynamics. Current opinion in biotechnology *19*, 404-409.

Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D.A., and White, F.M. (2007). Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. Proc Natl Acad Sci U S A *104*, 5860-5865.

Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H.D., Grantcharova, V., Lauffenburger, D.A., and White, F.M. (2006). Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. Mol Syst Biol *2*, 54.

Wright, C., Nicholson, S., Angus, B., Sainsbury, J.R., Farndon, J., Cairns, J., Harris, A.L., and Horne, C.H. (1992). Relationship between c-erbB-2 protein product expression and response to endocrine therapy in advanced breast cancer. British journal of cancer *65*, 118-121.

Wu, C.C., MacCoss, M.J., Howell, K.E., Matthews, D.E., and Yates, J.R., 3rd (2004). Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. Anal Chem *76*, 4951-4959.

Yaffe, M.B., and White, F.M. (2007). Signaling netwErks get the global treatment. Genome Biol *8*, 202.

Yarden, Y., and Sliwkowski, M.X. (2001). Untangling the ErbB signalling network. Nat Rev Mol Cell Biol *2*, 127-137.

Zhang, Y., Wolf-Yadlin, A., Ross, P.L., Pappin, D.J., Rush, J., Lauffenburger, D.A., and White, F.M. (2005). Time-resolved mass spectrometry of tyrosine

phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics *4*, 1240-1250.

Zubarev, R.A., Horn, D.M., Fridriksson, E.K., Kelleher, N.L., Kruger, N.A., Lewis, M.A., Carpenter, B.K., and McLafferty, F.W. (2000). Electron capture dissociation for structural characterization of multiply charged protein cations. Anal Chem *72*, 563-573.

II. Identification of therapeutic targets to revert Tamoxifen resistance in breast cancer by quantitative proteomic analyses of signaling network

II.1 SUMMARY

Estrogen and estrogen receptor biology have been extensively studied in the past fifty years, resulting in significant contributions to the development of breast cancer therapeutics. Tamoxifen resistance is an unmet medical need affecting a large number of patients with ER-positive breast cancer that have taken therapeutic selective estrogen receptor modulators (SERMs). Non-genomic functions of estrogen receptor (ER) provided many insights into the modulation of signaling pathways by ER and potential mechanisms of Tamoxifen resistance. While the non-genomic functions of ER affect the activities of kinases at various nodes, a comprehensive map of signaling networks influenced by ER and changes within these networks in disease settings such as Tamoxifen resistance are still lacking. By plotting a global map of such a signaling network and identifying differences between Tamoxifen sensitive and Tamoxifen resistant cell lines, one would be better able to propose novel therapeutic candidates.

In this chapter, I describe the development of Tamoxifen resistant breast cancer cell lines from a Tamoxifen sensitive cell line. The Tamoxifen sensitive

model is a parental MCF7 cell line, which is a classic model in studies of estrogen biology. Two Tamoxifen resistant cell lines are described in this chapter. One of the two Tamoxifen resistant cell lines is a HER2 overexpressing derivative of parental MCF7. The second Tamoxifen resistant cell line is a derivative of parental MCF7 cells which parental MCF7 were exposed to Tamoxifen for six months in order for cells to acquire Tamoxifen resistance.

Next, I describe the designs and results of quantitative phosphoproteomic analyses of tyrosine phosphorylation events, mainly focusing on signal transduction downstream of HER2 tyrosine kinase and Src kinase. These analyses resulted in the identification and quantification of 120 phosphorylation sites. MS analyses highlighted several important differences in cell-specific aspects of Tamoxifen treatment between Tamoxifen sensitive and Tamoxifen resistance cell line models. For instance, network diagrams of each cell line showed upregulation of PI3K-Akt pathways, MEK-ERK1/2 pathways, and Src/FAK/Abl pathways in Tamoxifen resistant cells. In addition to MCF7 cell line models, quantitative phosphoproteomic analyses were extended to human primary breast tumor samples. Two tumor samples were compared to study similarities and differences

between in vivo and in vitro analyses. A tumor from a Tamoxifen treated patient who suffered from recurrence was the first sample, and a tumor from a patient before Tamoxifen treatment was the second sample. I found similarities between data obtained from cell line analyses and primary tumors in that ERK1/2 and Src phosphorylation sites increased their intensities in response to Tamoxifen.

In the last chapter, I describe phenotypic assays to illustrate how quantitative phosphoproteomic analyses lead to the identification of effective therapeutic candidates. We employed small molecule inhibitors, some of which are FDA-approved, to validate our hypotheses developed from the MS analyses. We hypothesized that PI3K-Akt pathways, MEK-ERK1/2 pathways, and Src/FAK/Abl pathways are upregulated, play important roles, and are potential therapeutic target in Tamoxifen resistant cells. Inhibition of these pathways decreased the number of viable Tamoxifen resistant cells. The outcome of inhibition experiments showed cell-context specific responses to each compound treatment, with Src/Abl inhibition being the most effective treatment against both Tamoxifen resistant cell line models. In addition, phosphoproteomic analyses from a collaborative project on NOS inhibitor treatment, migration and invasion assays,

and preliminary results from experiments indicating potential IGF1-R/HER2 cross-talk are described.

This project was designed by Hideshiro Saito-Benz, Steven R. Tannenbaum, and Forest M. White. The experimental procedure was conducted by Hideshiro Saito-Benz, Kertikeya Pant, Kristen Naegle, Yi Wang, and Rachel Niehus.

II.2 MATERIALS and METHODS

II.2.1 Cell Culture, Cell line derivation.

MCF7 cell line is a breast epithelial adenocarcinoma cell line that expresses ER, is responsive to estrogen for its growth, and has been a classical model to study anti-estrogen treatment and resistance mechanisms. MCF7 were obtained from American Type Culture Collection (ATCC) and all Tamoxifen resistant derivatives were established in the White laboratory, with an exception of MCF7-HER2 which was a kind gift from Dr. Jack Buesmans (Astra Zeneca). MCF7 and their derivatives were cultured in the full growth media consisting of phenol red-free DMEM with 10% FBS and incubated in 95% air/5% CO_2 atmosphere at 37°C. MCF-TAM cell line was generated as previously described (Knowlden et al., 2003). In brief, parental MCF7 cells were cultured in phenol red-free DMEM with 10% charcoal-stripped FBS and 100nM 4-hydroxyltamoxifen (4-OHT, Sigma) for six months. Approximately for the first four weeks, cell growth remained static. As cells resumed slow growth after four weeks, cells were passed

at 1:6 dilutions at 70% confluence for total six months. MCF7-Flp cell lines are developed by stably transfecting parental MCF7 with Flp-In plasmids included in the Flp-Kit (Invitrogen).

II.2.2 Drug treatment for Viability Assay.

For cell viability assay, cells ($1.0x10^4/cm^2$) were seeded on six well plates in triplicate for 24 hours before compound treatments. At each time point of measurement, adherent cells were trypsinized and cell densities were counted on ViCell Counter (Beckman Coulter) according to manufacturer's manual. The media with compounds were replenished every three days. PI3K inhibitors (LY294002 and PI103) were purchased from CalBiochem, geldenamycin and 17AAG were purchased from Alexis Biochemical, and dasatinib was purchased from LC Laboratories.

II.2.3 Peptide sample preparation, , and quantitative mass spectrometry.

For mass spectrometry analyses, all cells were seeded in the full growth media at $1.0x10^4/cm^2$ on 15cm plates in biological duplicates. After 24 hours, media was switched to the full growth media with 100nM 4-OHT or vehicle control. Media and small molecule inhibitors were replenished every three days until cell lyses. On the appropriate day, cells were lysed, processed and isotopically labeled as previously described (Zhang et al., 2005). In brief, cells were lysed in 8M urea buffer with 100 mM sodium orthovanadate and proteins in cell lysate were reduced with DTT and alkylated with iodoacetamide to block free thiols on cysteine side chains. Samples were subjected to tryptic digestion and the resulting peptides were desalted on a C18 reverse phase cartridge (Millipore). Peptides were eluted with 10mL of 25% acetonitrile, 0.1% acetic acid, and aliquoted in 500μg fractions prior to lyophilization and storage at -80C until time of analysis.

II.2.4 iTRAQ labeling of peptides

Six samples were paired into two MS analyses combinations, each consisting of four samples that were covalently reacted with 4 isotopic labels of

iTRAQ. The first sample combination was labeled with iTRAQ reagent as follows: iTRAQ 114: parental MCF7 with vehicle treatment, iTRAQ 115: parental with 15 day 4-OHT treatment, iTRAQ 116: MCF7-HER2 with vehicle treatment, and iTRAQ117: MCF7-HER2 with 15 day 4-OHT treatment. The second sample combination was labeled with iTRAQ reagent as follows: iTRAQ 114: MCF7-TAM with vehicle treatment, iTRAQ 115: MCF7-TAM with 15 day 4-OHT treatment, iTRAQ 116: MCF7-HER2 with vehicle treatment, and iTRAQ117: MCF7-HER2 with 15 day 4-OHT treatment.

II.2.5 Peptide Immunoprecipitation

After iTRAQ labeling, peptides were combined and subjected to peptide immunoprecipitation as described (Zhang et al., 2005) with the following exceptions: iTRAQ labeled peptides were incubated with 12µg of antiphosphotyrosine antibody (P-Tyr-100, (Cell Signaling Technology)) and 12µg of antiphosphorylation antibody 4G10 (Millipore) in 200ug of immunoprecipitation buffer for 8 hours at 4C, followed by incubation with 10ug of protein G Plus-agarose beads (Calbiochem) overnight.

II2.6 Immobilized metal affinity chromatography (IMAC) and Mass Spectrometry

Following elution from peptide immunoprecipitation, tyrosine phosphorylated peptides were further enriched by immobilized metal affinity chromatography (IMAC) to remove non-specific binding peptides resulting from peptide immunoprecipitation (Moser and White, 2006). Peptides retained on the IMAC column were eluted with 250mM sodium phosphate at pH8.0 and analyzed by electrospray ionization liquid chromatography tandem MS on QSTAR XL Pro Mass Spectrometer (Applied Biosystems) (Zhang et al., 2005). MS/MS spectra were extracted, searched with Mascot (Matrix Science), and quantified with ProQuant (Applied Biosystems). Phosphorylation sites and peptide sequence assignments were validated by manual confirmation by assigning spectra of raw MS/MS data. Area under the peak of iTRAQ reporter ions from phosphopeptides were normalized with values from iTRAQ reporter ion peak areas of non-phosphorylated peptides from high expression level proteins found in the

supernatant of the immunoprecipitation and corresponding to total peptide amount for each condition. Each condition was normalized against the MCF7-HER2 cell line with vehicle treatment to obtain relative percentages across all six conditions.

II.2.7 Preparation of Human Tumor Samples

Two human primary breast tumor samples (recurrent, or before treatment) are received from Dr. Morag Park at McGill University in Canada. Each tumor was approximately 100mg with pink color. Approximately 20mg of each tumor is cut by surgical razor blades and remaining tumors were stored in -80C. The 20mg of tumors were suspended in 1ml of 8M urea, homogenized on ice, and processed for tryptic digest in the same way as cell line lysate samples were processed. The digested sample was passed through the fat-removal membrane (Waters) before passed through SepPak C18 column. Recurrent tumor was labled with iTRAQ114, and 116, and before-treatment tumor sample was labeled with iTRAQ115 and 117 to obtain standard errors and averages for every phpshorylation sites identified from a single MS analysis.

II.2.8 ELISA measurement

ELISA measurements were conducted according to manufacturer's protocol (Invitrogen). In brief, cells were seeded at $1.0x10^4/cm^2$ on 10cm plate in biological duplicates. After 24 hours, media was changed to the full growth media with 100nM 4-OHT or vehicle control. Cells were harvested at 70% confluence for vehicle control, or after 15 days 4-OHT treatments, with denaturing cell extraction buffer (Invitrogen) as recommended by manufacturer's protocols. Either 50µg, 25µg, or 10µg of protein lysates were used for detection of phosphoproteins and total protein depending on abundance or proteins in each lysate in order to keep the ELISA readings within a linear range of detection.

II.2.9 Bioinformatic Analysis

I created an aligned phosphopeptide background by expanding the sequence around every site of phosphorylation to +/-7 amino acids using the Entrez

Protein database. This yielded 129 aligned singly phosphorylated peptides. We then created a 4-OHT sensitive MCF7-HER2 foreground by choosing those peptides that increased at least two-fold in the MCF7-HER2 cells in response to 4-OHT, which yielded 35 phosphopeptides and 42 singly phosphorylated aligned peptides. We searched for enriched motifs within +/-5 amino acids of the oriented phosphorylation site using a greedy search algorithm(Joughin et al., 2009). In order to examine submotifs off of a parent branch in this algorithm, we required at least 25% of the foreground matched the parent motif and with at least a significance of 0.01, as judged by the hypergeometric test. Motif analysis yielded two final motifs with a p-value less than 0.01. If we use Bonferroni correction to adjust these scores for multiple hypothesis testing, the only motif that remains significant is y.[E|D]P with a p-value of 0.0058 and a corrected p-value of 0.064. Given Bonferroni is an overly severe correction due to interdependence of the hypotheses tested, 0.064 is the upper bound on the significance of this motif, but the motif is in fact likely to be more significant than this bound. To determine the regulating kinase(s) responsible for enriched motifs we examined known linear amino acid sequences in HPRD (Keshava Prasad et al., 2009) and utilized Scansite

kinase predictions (Obenauer et al., 2003).

II.2.10 NOS inhibitor treatments

Tamoxifen resistant cells (MCF7-HER2 and MCF7-TAM) were treated with 1) 100nM Tamoxifen alone, 2) 100nM Tamoxifen and 1mM 1400W NOS inhibitor, 3) 100nM Tamoxifen and 1mM S-methylthiocitrulline (SMTC), and 4) 100nM Tamoxifen and 0.3mM SMTC for 15 days. Cell culture media and compounds were replenished every 3 days.

II.2.11 Growth factor treatments

Cells were seeded at $1.0x10^4/cm^2$ on 10cm plate in biological duplicates. After 24 hours, media was changed to the full growth media with vehicle control, 100nM 4-OHT, with or without growth factors. EGF, HRG, IGF1 were dosed at 100ng/ml. Media was replenished every three days. Parental MCF7 cells were treated for nine days before counting viable cell numbers. MCF7-HER2 cells were

treated for six days before counting viable cell numbers.

II.2.12 Migration assay

All cells were seeded in the full growth media at $1.0x10^4/cm^2$ on 15cm plates in biological duplicates. After 24 hours, media was switched to the full growth media with 100nM 4-OHT or vehicle control and cultured for either six, nine or fifteen days as the same conditions as mass spec samples were prepared. Media and small molecule inhibitors were replenished every three days until cell samples were tested for migration ability by Boyden chamber migration assay. Parental MCF7, MCF7-HER2, and MCF7-TAM are grown for nine, six and nine days respectively with vehicle treatment and seeded onto Boyden chambers. Tamoxifen treated cells are grown for 15 days before seeded onto Boyden chambers. Migration assay was conducted following the White lab migration assay protocol. In brief, cells were directly seeded at $1.0x10^5/cm^2$ in 100ul into surface of Boyden chamber. Lower chamber was filled with the full growth media with 10% FBS. Cells were undisturbed in cell culture incubator for 24, 48, and 72

hours, and stained with crystal violet cell stain solution. Cells on a top layer of membrane were removed by cotton swabs, and numbers of cells on bottom layer of Boyden chamber membrane were counted under a microscope. Three analytical replicate wells were counted and experiments were repeated twice.

II.2.13 Invasion Assay

All cells were seeded in the full growth media at $1.0x10^4/cm^2$ on 15cm plates in biological duplicates. After 24 hours, media was switched to the full growth media with 100nM 4-OHT or vehicle control and cultured either six, nine or fifteen days for same conditions as MS samples were prepared. Media and small molecule inhibitors were replenished every three days until cell samples were tested for migration ability by Boyden chambers filled with matrigels. Parental MCF7, MCF7-HER2, and MCF7-TAM are grown for nine, six and nine days respectively with vehicle treatment and seeded onto Boyden chambers. Similarly, Tamoxifen treated cells are grown for 15 days before seeded onto Boyden chambers. Invasion assay was conducted following the White lab migration assay

protocol. In brief, matrigel was diluted at 30 fold in serum free growth media, and 100ul diluted matrigel was spread on Boyden chamber prior to UV light exposure. Cells were seeded at $1.0x10^5/cm^2$ in 100ul into surface of Boyden chamber filled with sterile matrigel. Cells were undisturbed in cell culture incubator for 24, 48, and 72 hours, and stained with crystalviolet cell stain solution. Matrigel and cells on a top layer of membrane were removed by cotton swabs, and numbers of cells on bottom layer of Boyden chamber membrane were counted under a microscope. Three analytical replicate wells were counted and experiments were repeated twice.

II.3 RESULTS

II.3.1 Development of acquired Tamoxifen resistant MCF7 (MCF7-TAM).

In order to obtain MCF7 cells that naturally developed resistance to Tamoxifen, parental MCF7 cells had been continuously exposed to 100nM 4-hydroxyTamoxifen for six months in low estrogen media following method described in Knowlden et al (Knowlden et al., 2003). This approach mimics clinical conditions when patients are put on adjuvant Tamoxifen therapy after surgical removal of primary tumor and develops Tamoxifen resistance or Tamoxifen as prophylaxis for women with high risk of breast cancer who develops Tamoxifen resistant breast tumors. Cells were seeded such that 100 mm plates became approximately 25% confluent, and media were switched to DMEM with 10% CSFBS supplemented with 100nM 4-hydroxyTamoxifen (4-OHT). Growth of cells slowed down and resulted in growth arrest conditions for the first four weeks. During this period, a significant number of cells died, presumably due to combinations of 100nM 4-hydroxyTamoxifen and low estrogen in the media. The growth of

remaining cell population stayed static over the first month. After four weeks, a growth rate gradually increased to a noticeable level and cells on a 100mm plate reached to confluence at the end of third month. From this point cells on 100 mm plates were continuously passaged in a same treatment for additional three months. There were two noticeable morphological changes at early stages of Tamoxifen treatment after the second month and late stage of Tamoxifen treatment at the sixth month. After two months of exposure to 100nM Tamoxifen, cells showed broadened morphology that was distinct from parental MCF7 (Figure 1).

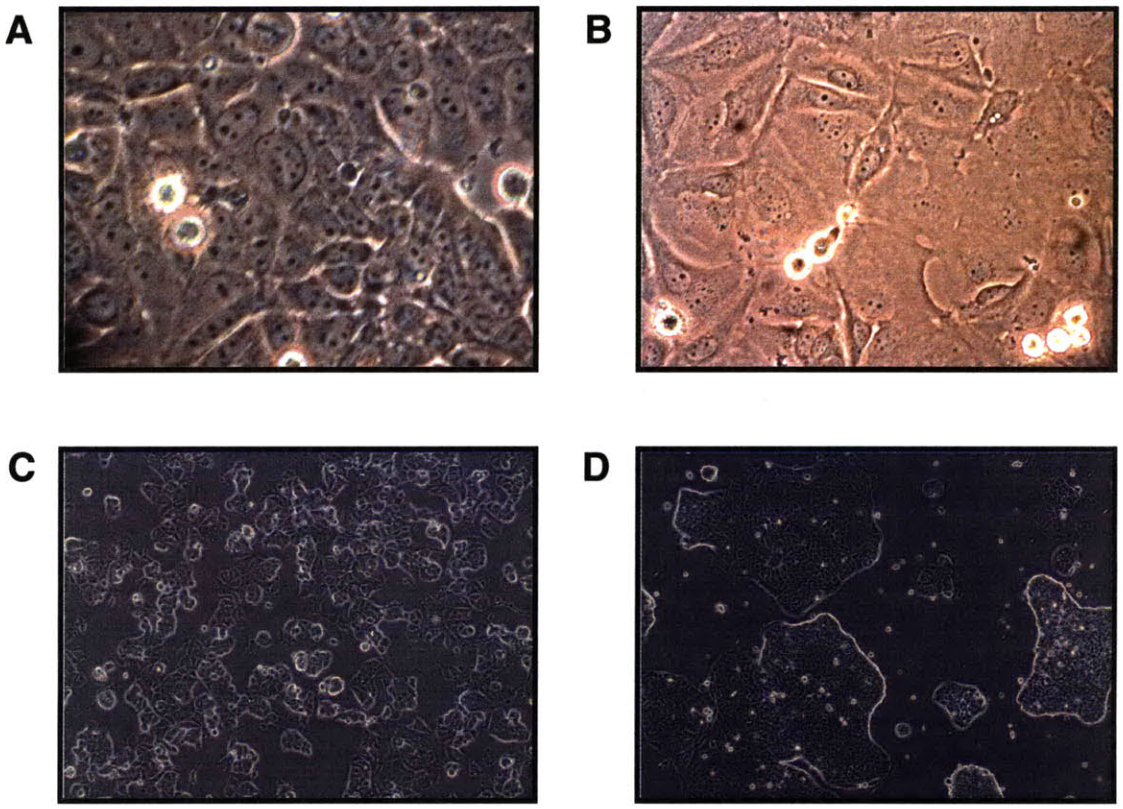


Figure 1. Images of parental MCF7 and MCF7-TAM under microscope. A). parental MCF7 under confluent condition. 400x magnification. B) MCF7 cells exposed to Tamoxifen for 2 months shows spread out cellular phenotype. 400x magnification. C). parental MCF7 under confluent condition. 100x magnification. D) MCF7 cells exposed to Tamoxifen for 6 months (MCF7-TAM) grows with tight cellular density with close contact to each other. 100x magnification.

This broadened morphology was indicative of stressed cells presumably due to low estrogen and presence of Tamoxifen. At the fifth and sixth months of low Tamoxifen treatment, cells started to grow noticeably faster compared to the third

month, and also showed distinct morphology compared to that of second-third months or that of parental MCF7 (Figure 1D). The cells grew in high area density, and area occupied by a single cell became visually less than parental MCF7 cells. When cell diameter observed during trypsinization, cell size did showed less than 10% changes in cell diameter (approximately 19μm (MCF7-TAM, six months) compared to 20μm (parental MCF7) when counted Vi-Cell XR). Tightly packed cells indicates that cell-cell contact proteins have been altered in order for cells to tightly contact to each other. Although long-term treatment of parental MCF7 cells with 100nM 4-OHT has previously been shown to result in increased EGFR/HER2 expression (Knowlden et al., 2003), MCF7-TAM cells generated in this study did not demonstrate noticeable increase in HER2 expression (Figure 2A), and did not decrease estrogen receptor (ER) expression (Figure 2B)

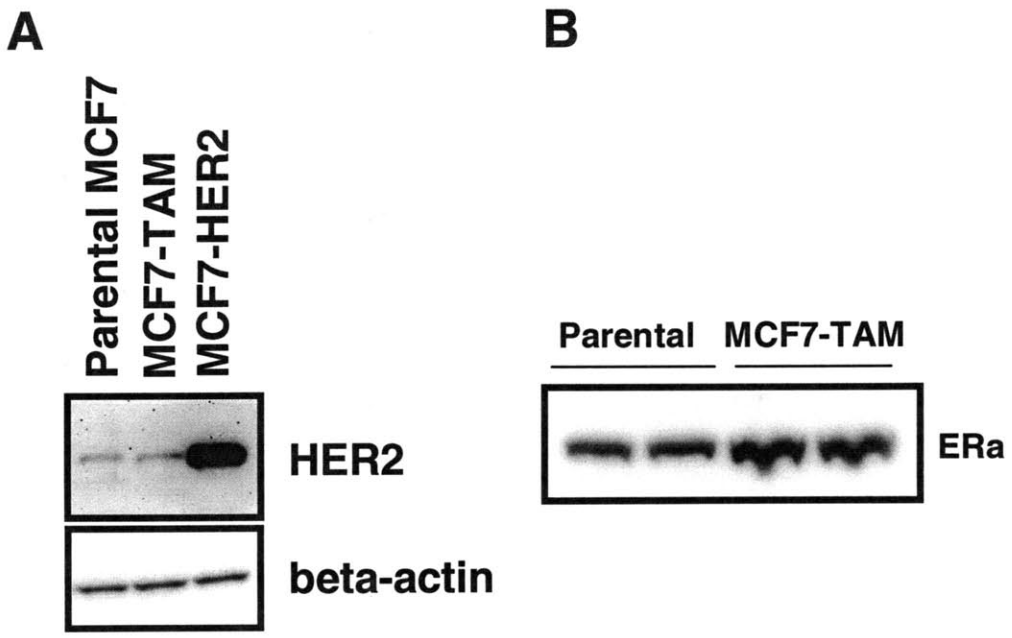


Figure 2. Western blot for HER2 and ER. (A) Total protein expression level of HER2 in parental MCF7, MCF7-TAM and MCF7-HER2 was compared by western blot. parental MCF7 and MCF7-TAM expresses similar level of HER2, while MCF7-HER2 express approximately 40 fold higher level of total HER2 protein. β-actin was used to control for total protein loading. (B) Total protein expression level of ER in parental MCF7 and MCF7-TAM. MCF7-TAM cells maintain ER expression after 6 months of exposure to 100nM 4-OHT.

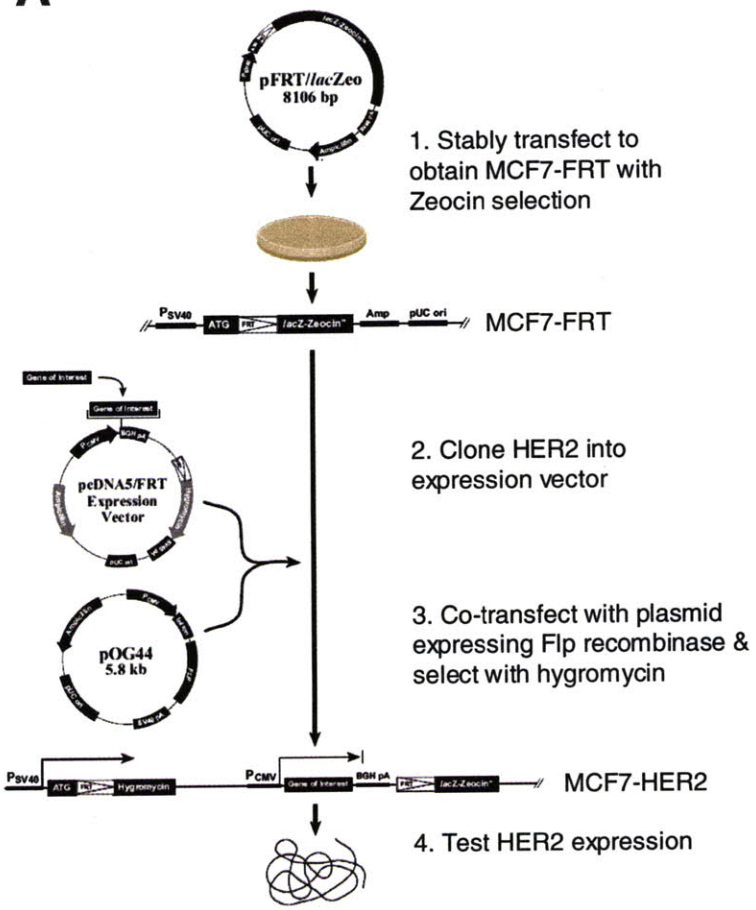
II.3.2 Obtaining HER2 overexpressing MCF7

HER2 overexpressing MCF7 has been previously documented by several groups to become Tamoxifen resistant in vivo (Benz et al., 1992; Shou et al., 2004).

While both Forest and I contacted the authors of the paper, we could not receive the cell line clone. The reason was mainly due to the patent issue that restricted the distribution of the MCF7-HER2 clone 18 described Benz et. al and Show et al(Benz et al., 1992; Shou et al., 2004). The MCF7-HER2 clone 18 was patented under Genentech and was restricted from distribution to both academic and industrial labs. Because of inability to access the published MCF7 clone that overexpress HER2, we decided to produce our own MCF7 cell line with HER2 overexpression

The Flp-In stable transfection system is the stable transfection methodology which allows integration and expression of gene of interest in mammalian cells at a specific genomic location (Figure 3).

A



B

Advantages:

1. Once MCF7-FRT is produced, second transfection is mediated by Flp recombinase rather than relying on random integration event
2. The integration event occurs at same location in the genome, producing isogenic clones with same expression level
3. Production of stable clone with different gene expressing same level is feasible, allowing direct comparison of effect of two genes
4. Production of multiple different clones with gene of interests at preferred expression level is possible

Disadvantages

1. Must produce MCF7-FRT with classical stable transfection method as Flp-In host MCF7-FRT is not commercially available from Invitrogen
2. Selection with Zeocin is not as easy as as other selection agent such as G418
3. Takes longer time than retroviral transfection followed by FACS sorting based on reporter expression

Figure 3. Flp-In stable transfection system A) Overall scheme of transfection steps in Flp-In stable transfection system. B) Advantages and disadvantages compared to traditional stable transfection or retroviral transfection.

It was one of the latest stable-transfection methods that Invitrogen was advertising around the time when we decided to produce MCF7-HER2. Amy Nicholes was producing HER2 overexpressing LNCaP Prostate cell lines using

Flp-In system in the White Lab at the time I was going to make MCF7-HER2. The system involves introduction of single or multiple Flp recombinase target (FRT) site(s) into the genome of the parental MCF7 cells to create Flp-In ready host MCF7-FRT cell line. An expression vector containing HER2 was then integrated into the genome of the MCF7-FRT at the specific FRT location via Flp recombinase-mediated DNA recombination (O'Gorman et al., 1991). There are multiple advantages of the Flp-In stable transfection system compared to traditional stable transfection or by retroviral transfection. The advantages and disadvantages are summarized in Figure 3B. Once the Flp-In host cell line containing an integrated FRT site has been created, subsequent generation of Flp-In cell lines expressing the genes of interest is rapid and efficient because integration is directed by Flp recombinase rather than random integration event. In addition, the system allows the generation of an isogenic stable cell line as the gene of interests is integrated into the same site in the genome. This is useful when effects of two genes are needed to be compared. For example, if effects of HER2 and constitutively active spliced variant HER2 are assessed, it is desirable that expression levels are kept at the same level. It is challenging to obtain such

clones by conventional methodologies, but theoretically Flp-In system allows to produce two MCF7 derivatives that has integrated different genes at a same part of host genome to derive the expression at same levels.

Successful stable integration of the linear plasmid and expression is random and rare, which is estimated to occur in less than one in a million transfected cells. Therefore optimizing the transfection condition was important before starting the selection process with antibiotics. Similarly, as excessively high antibiotic concentration can select out successfully transfected cells, determining the lowest antibiotic concentration possible for selection was important so that all untransfected cells are killed with antibiotics while only successfully transfected cells survives. To identify a optimal condition, transfection with Lipofectamine2000 (Invitrogen) was tested with various ratio of Lipofectamine2000 to linearized plasmid DNA to maximize a transfection efficiency while a minimizing toxicity. Among Lipofectamine:DNA ratios tested between 1:0.5 and 1:5, the 1:2 ratio with 30μg plasmid DNA was shown to yield maximal transfection efficiency with minimal toxicity based on GFP expression of viable cells. To determine optimal antibiotic concentration, MCF7 cells transfected with an empty vector was subjected to

selection with incremental concentration of Zeocin antibiotics from 10μg/ml to 800μg/ml (Invitrogen). An optimal Zeocin concentration to select MCF7 transfected with GFP vector was determined to be 25μg/ml, and this concentration took approximately 10-12 days to kill approximately 80% of MCF7 transfected with a control vector. To transfect the MCF7 with linearized pFRT/*lacZeo* 100 mm plate was brought to 50% confluent, and cells were transfected with Lipofectamine2000 at 1:2 ratio with 30ug linearized pFRT/*lacZeo* plasmid or control vector pcDNA/GFP. 24 hours after the transfection, successful transfection was confirmed by GFP expression in control cells. At this point, >70% confluent 100 mm plates were subcultured at 1:4 into four 150 mm plates and cells were allowed to attach for 24 hours before the start of antibiotic selection. The resulting plates were approximately 20% confluent and selection with 25μg/ml Zeocin was started at 48 hours after the transfection. The media with 25μg/ml Zeocin was replenished every two days to remove dead cells and to sustain antibiotic pressure. After 10 days of antibiotic selection, significantly more than expected number of cells looked viable under microscope; this was most likely due to incomplete selection with 25μg/ml. Therefore, Zeocin concentration was increased to 50μg/ml for all 150

mm plates. However, significant changes were not observed in the next five days, and therefore Zeocin concentration was increased again. Zeocin concentration was increased to 100μg/ml in one third of the plates, and to 200μg/ml in another one third of the plates, and to 400μg/ml in the last one third of the plate. After another one week of selection, presence of viable cell colonies became clear among almost empty plates (Figure 4A).

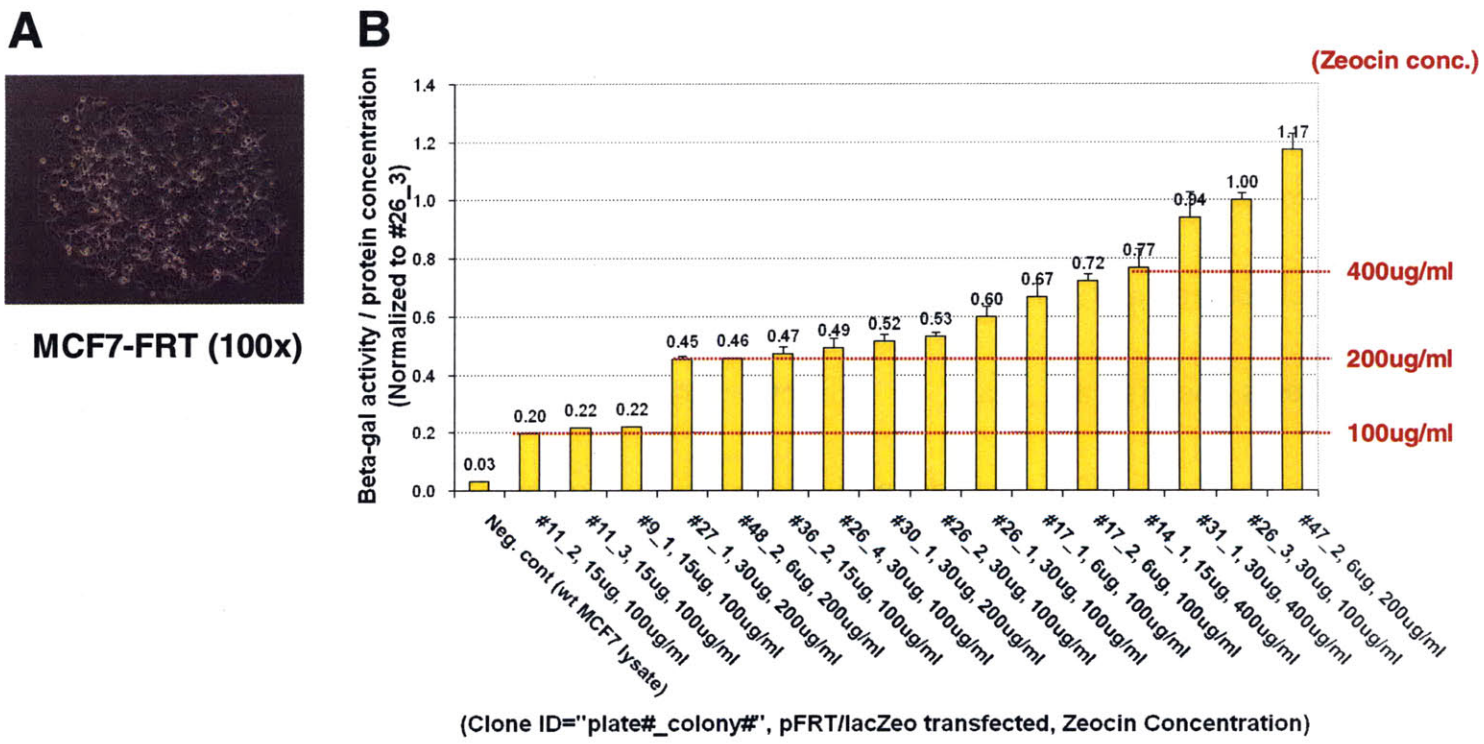


Figure 4. MCF7-FRT clones (A) Viable MCF7-FRT colony under microscope after Zeocin selection. 100x magnification (B) β-gal activity per protein concentration of each clone. Note the higher the β-galactosidase activity, the higher the Zeocin selection the colony can survive.

The colonies were marked, isolated by clone isolation cylinder, and each colony was transferred into 24 well plates to expand. From this point, cells in each well were considered as individual clones. After reaching over 70% confluence, clones in 24 wells were trypsinized, and all cells were transferred into 6 well plates to further expand. Once 6 wells became >70% confluent (equivalent to $2x10^6$ cells) the cells were trypsinized and transferred into one 100 mm plate. Once a 100mm plate were >70% confluent (15×10^6 cells), cells were subcultured at 1:3 into three 100 mm plates. Once three 100 mm plates were >70% confluent, cells from one of three 100 mm plates were collected for the assay to measure β-galactosidase activity resulting from stably integrated *lacZeo* gene, and cells from two other plates were frozen down to save the clones. The protein concentration from the each sample was also measured with a BCA assay. To determine the expression level of β-galactosidase from stably integrated *lacZeo* gene (which will proportionally correlate with HER2 expression level after HER2 integration), β-galactosidase activity from all clones was determined and normalized to MCF7-FRT clone ID #26_3 which was the first clone isolated. B-galactosidase activity per protein

concentration of all clones is shown in Figure 4B. Notably, β-galactosidase activity level had approximately six fold difference between clones, and the Zeocin selection concentration proportionally correlated with expression level. In the next step, several selected clones will be co-transfected with pcDNA5/FRT/HER2 and pOG44, and subjected for hygromycin selection to obtain MCF7-HER2 clones. Once MCF7-HER2 clones are obtained, relative HER2 expression level are measured by western blotting and MCF7-FLP Clone 6 (Med) and Clone 16 (High) was successfully transfected to overexpress HER2, and expression level was 5x and 10x higher than the parental MCF7 (Figure 5).

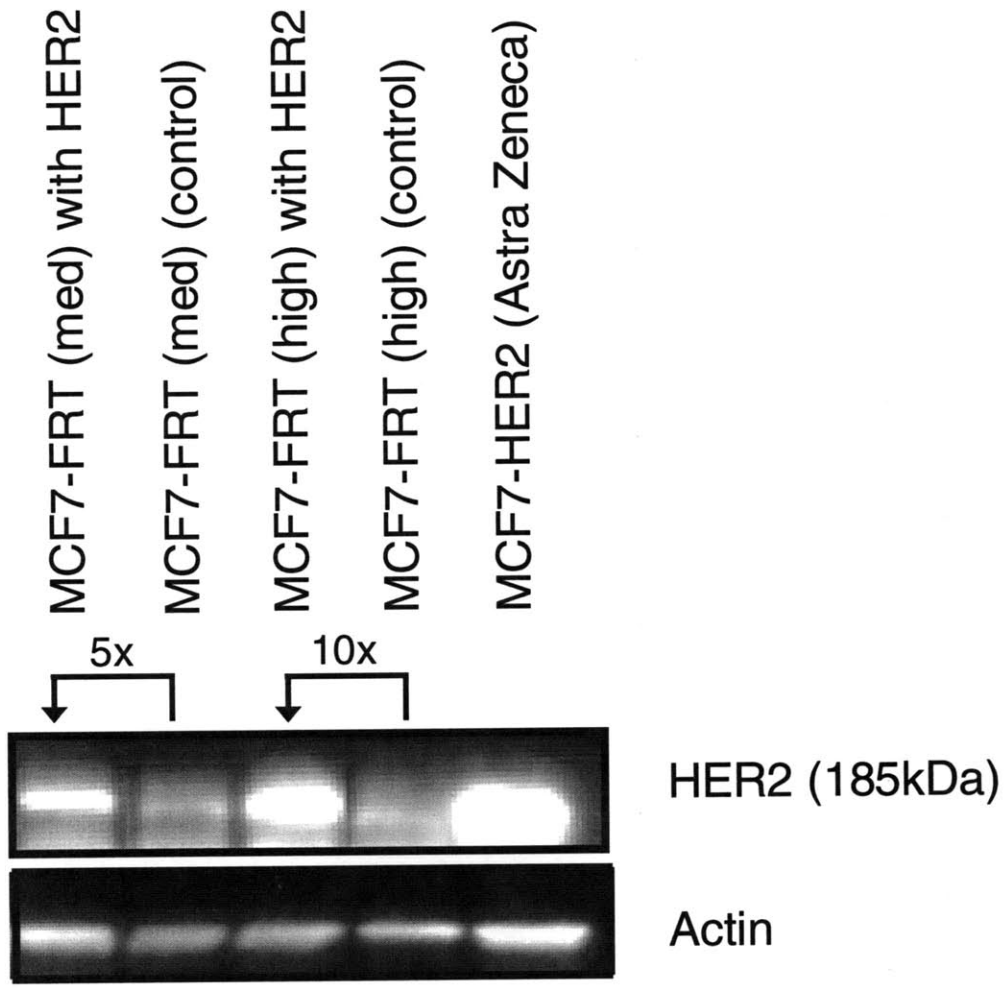


Figure 5. HER2 expression level of MCF7-HER2 via Flp-In relative to MCF7-HER2 from Astra Zeneca. MCF7-FRT (med) with HER2 and MCF7-FRT (high) with HER2 express 5x and 10x higher before HER2 transfection via Flp-In. MCF7-HER2 from Astra Zeneca overexpress HER2 at much higher level (30-40x).

Advantages of MCF7-FLP clones with different expression level of reporter gene expression is that transfection of other genes such as EGFR, IGFR-1 could be studied at desired expression levels of each genes.

While we derived HER2 overexpressing MCF7 via Flp-In system, we found an access to HER2 overexpressing MCF7-HER2 from Astra Zeneca. We got access to HER2 overexpressing MCF7 during the trip that Forest, Doug, Neil, Ale, and Hide made trip to Astra Zeneca in July 2007. I was talking to one of the researchers from the oncology group about my project, and they indicated that the group internally obtained HER2 overexpressing MCF7 from retrovirally transfecting the HER2 and selecting high overexpressing clones. MCF7-HER2 from Astra Zeneca was a clone of MCF7 cells with 30-40x overexpression of HER2 based on FACS selection with GFP as selection marker. Expression level of HER2 was validated in the White lab by western blots, and results showed that HER2 expression was approximately 40 times higher than parental MCF7 (Figure 2A). Because the original reports from Christopher Benz (Benz et al., 1992) used MCF7-HER2 clone 18 which overespress HER2 at 42 times higher than parental MCF7, we decided to use MCF7-HER2 from Astra Zeneca as a cell line of choice for the project. MCF7-HER2 (med) and MCF7-HER2 (high) produced from Flp-In study was saved for latter study as they could be used for other experiments such as comparing signaling events in response to liner increase in kinaes expression

levels.

II.3.3 Growth Profiles of parental MCF7, MCF7-TAM, and MCF7-HER2.

Once all three cell lines are obtained, we decided to establish their responsiveness to Tamoxifen. To determine the Tamoxifen sensitivity of parental and derivative MCF7 cells, cell growth assays were performed in the absence and presence of 4-OHT. As indicated in Figure 6, in the absence of 4-OHT, MCF7-HER2 cells had significantly higher growth rate than parental MCF7 cells or MCF7-TAM cells, presumably due to the pro-proliferative activating signal from HER2, while under these same conditions parental MCF7 and MCF7-TAM grew at approximately the same rate relative to each other.

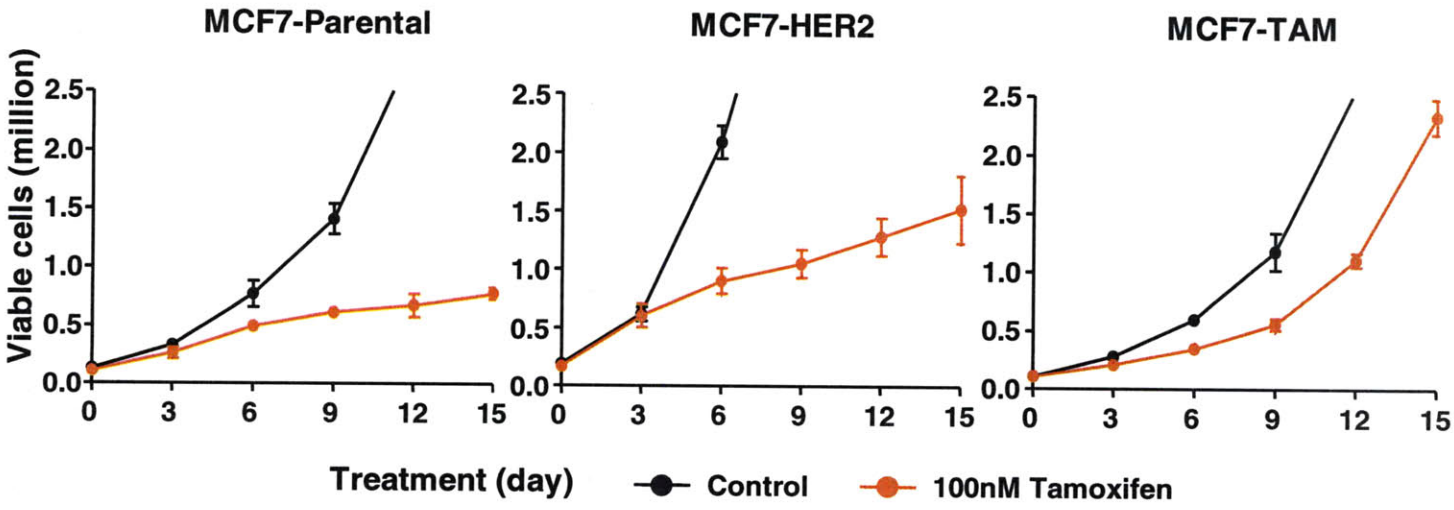


Figure 6. Growth profile of parental-MCF7, MCF7-TAM, and MCF7-HER2 in the presence or absence of 100nM 4-OHT treatment. Average viable cell number and standard deviation measured from three individual cell culture wells is plotted as described in the methods. The plot is representative of at least three independent biological replicates.

As expected, in the presence of 100nM 4-OHT parental MCF7 cells demonstrated a significantly decreased, almost static, growth rate, indicating a strong sensitivity to Tamoxifen for this cell line. In stark contrast to parental MCF7 cells, the MCF7-TAM cell line was minimally responsive to 4-OHT treatment and maintained a high growth rate in the presence of 4-OHT, indicating the almost complete resistance of these cells to this concentration of 4-OHT. MCF7-HER2 cell growth rates were significantly diminished by exposure to 100nM 4-OHT,

indicating a surprising sensitivity to Tamoxifen in this supposedly Tamoxifen resistant model system. Although exposure to 100nM 4-OHT led to a significant decrease in growth rate, MCF7-HER2 cells maintained a higher growth rate compared to parental MCF7 cells under the same condition, and thereby demonstrated an increased level of resistance relative to parental MCF7 cells. Tamoxifen resistance of the MCF7-TAM and MCF7-HER2 cells extended to a higher concentration of 4-OHT, as both cells demonstrated continued growth in the presence of 1μM 4-OHT, while parental MCF7 growth was completely abrogated under these same conditions (Figure 7).

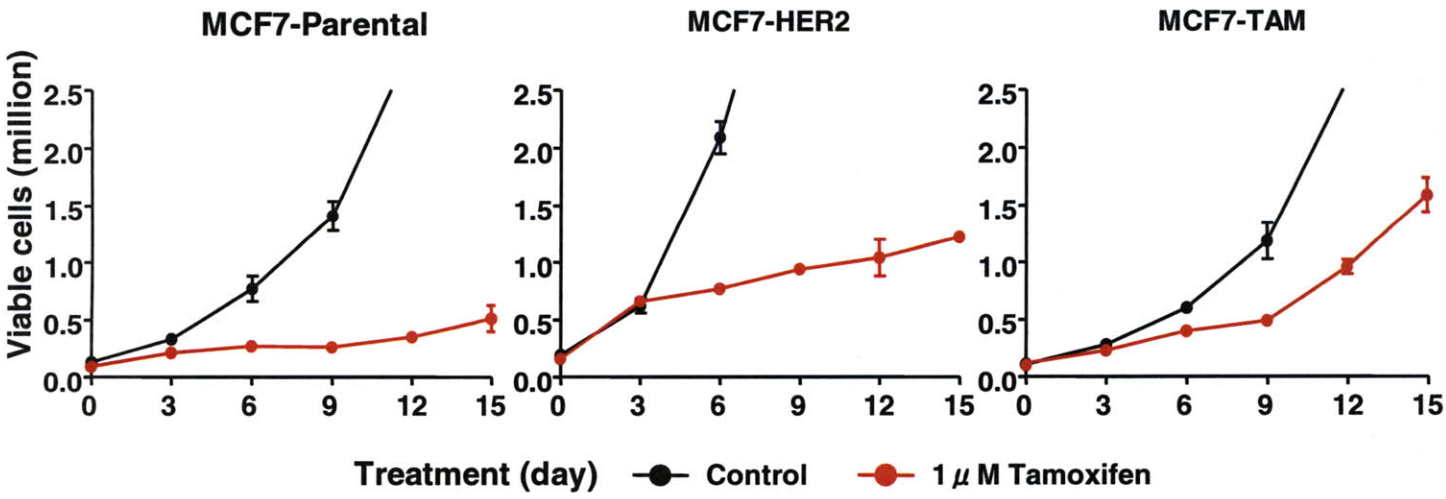


Figure 7. Growth profile of parental-MCF7, MCF7-TAM, and MCF7-HER2 in 1μM 4-OHT treatment. Similar to 100nM 4-OHT treatment, parental MCF7 is

growth suppressed beyond 15 days. MCF7-TAM is not significantly growth suppressed by 1μM 4-OHT treatment. MCF7-HER2 cells are affected by 4-OHT treatment, but growth rate continues to be positive, similar to that shown for 100nM 4-OHT exposure.

When three cell lines are grown in low-steroid media containing charcoal-stripped FBS, parental MCF7 failed to grow, MCF7-HER2 maintained slow grow rate, and MCF7-TAM grow most among all cell lines. These results indicates that HER2 overexpression supplements the growth signal to make MCF7 cells become less dependent on steroid media. On the other hand, MCF7-TAM cells were steroid-independent for it growth, and became self-sufficient in growth sustenance in steroid-depleted media condition. While MCF7-TAM maintaind comparable ER expression levels to the parental MCF7, it was possible that post-translational modification of the ER is significantly altered in MCF7-TAM such that ligand-independent activation of ER occurred in these population (Shou et al., 2004)

II3.4 Tyrosine phosphorylation signaling network analyses of Tamoxifen sensitive and resistant cells.

Although the estrogen receptor is best characterized as a nuclear receptor and transcriptional regulator, it has also been proposed that this receptor have a non-genomic role involving modulation of various signaling pathways (Bjornstrom and Sjoberg, 2005). Since 4-OHT binds to the ER and thereby inhibits ER activation, it is not surprising that this compound affects both estrogen-induced transcriptional regulation and the non-genomic component of ER activity. To date effects of 4-OHT on a ER non-genomic pathway has been shown only for selected proteins in several pathways (Pedram et al., 2002), and a systems-level view of 4-OHT effects on the signaling network is still lacking. To address this issue we have performed an unbiased, mass spectrometry-based, quantitative analysis (Zhang et al., 2005) of tyrosine phosphorylation mediated signaling network changes associated with 4-OHT treatment. Specifically, cell lysates were generated from cell culture plates that had either reached 70% confluence in the absence of 4-OHT, or had been treated with 100 nM 4-OHT for 15 days. As described in Figure 8, tryptic peptides obtained from each of the six cell lysates were isotopically labeled, with the MCF7-HER2 lysates serving as a common

normalization point for relative quantification across the six conditions.

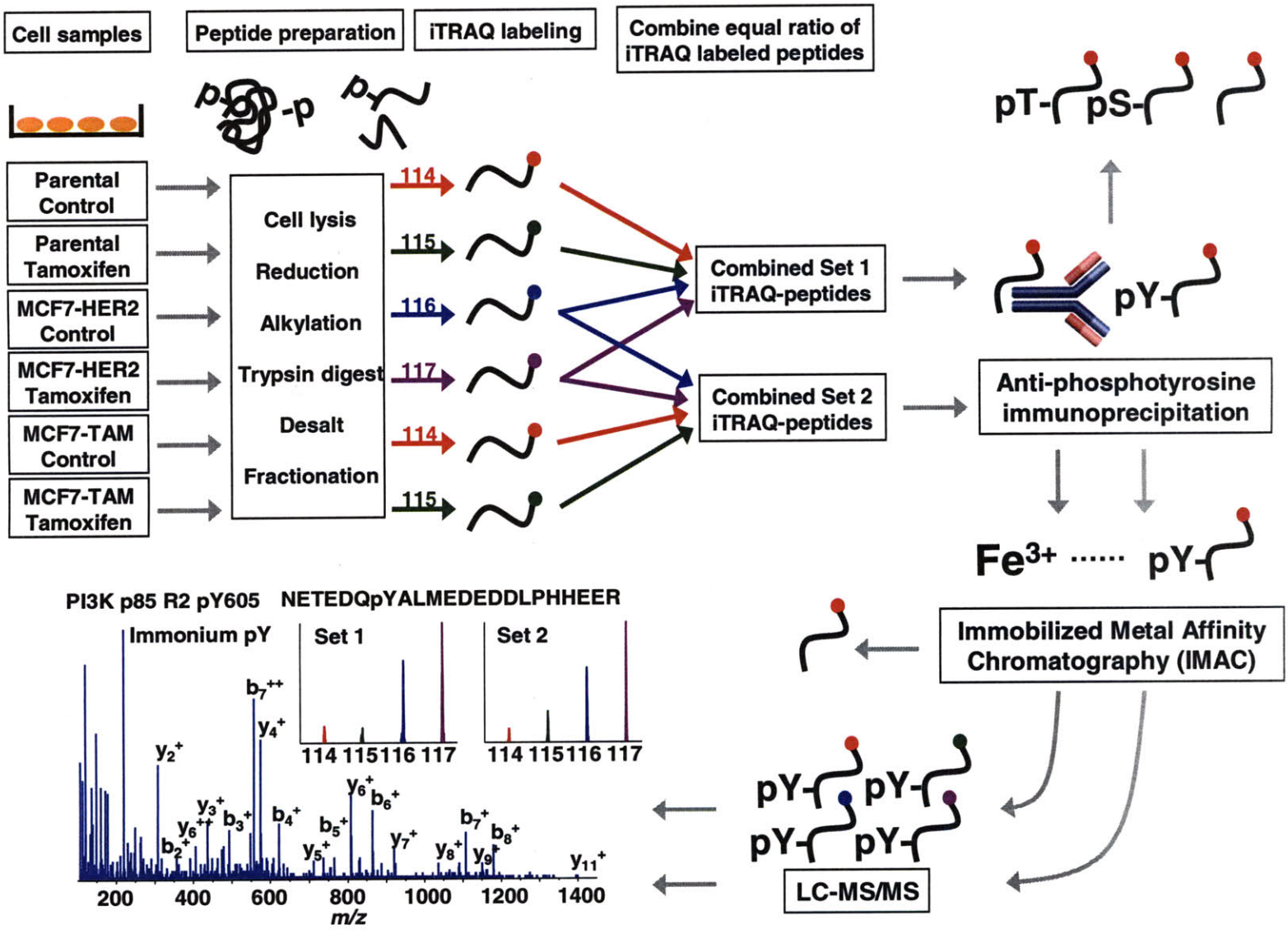


Figure 8. Schematic representation of the experimental approach for MS-based quantitative proteomic analysis of tyrosine phosphorylation. Following cell lysis and sample preparation, peptides from the six samples are labeled with iTRAQ for relative quantification. MCF7-HER2 cells are used to normalize the two experimental sets in order to obtain relative quantification across all six samples. A representative MS/MS spectrum and iTRAQ marker ion peaks for

pY605 on phosphoinositide-3 kinase p85β regulatory subunit 2 (PI3K R2 pY605) is shown.

Following stable isotope labeling, peptides from the appropriate samples were mixed and tyrosine phosphorylated peptides were isolated with a two-step enrichment protocol consisting of immunoprecipitation with a mixture of pan-specific anti-phosphotyrosine antibodies followed by enrichment of phosphorylated peptides by immobilized metal affinity chromatography (IMAC), as described previously (Zhang et al., 2005). Peptides retained on the IMAC column were eluted to a reverse-phase liquid chromatography column and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) on a quadrupole time-of-flight mass spectrometer. In total, 120 tyrosine phosphorylated peptides from 80 proteins were identified and quantified from analysis of biological replicates of the six samples (Table 1). Data-dependent peak selection for tandem MS fragmentation is known to introduce variability in the peptide identification and quantification in any given analysis; here this variance resulted in 80 of the 120 peptides being quantified across all conditions, while 15 and 25 were quantified for only four of the six conditions (Figure 9A).

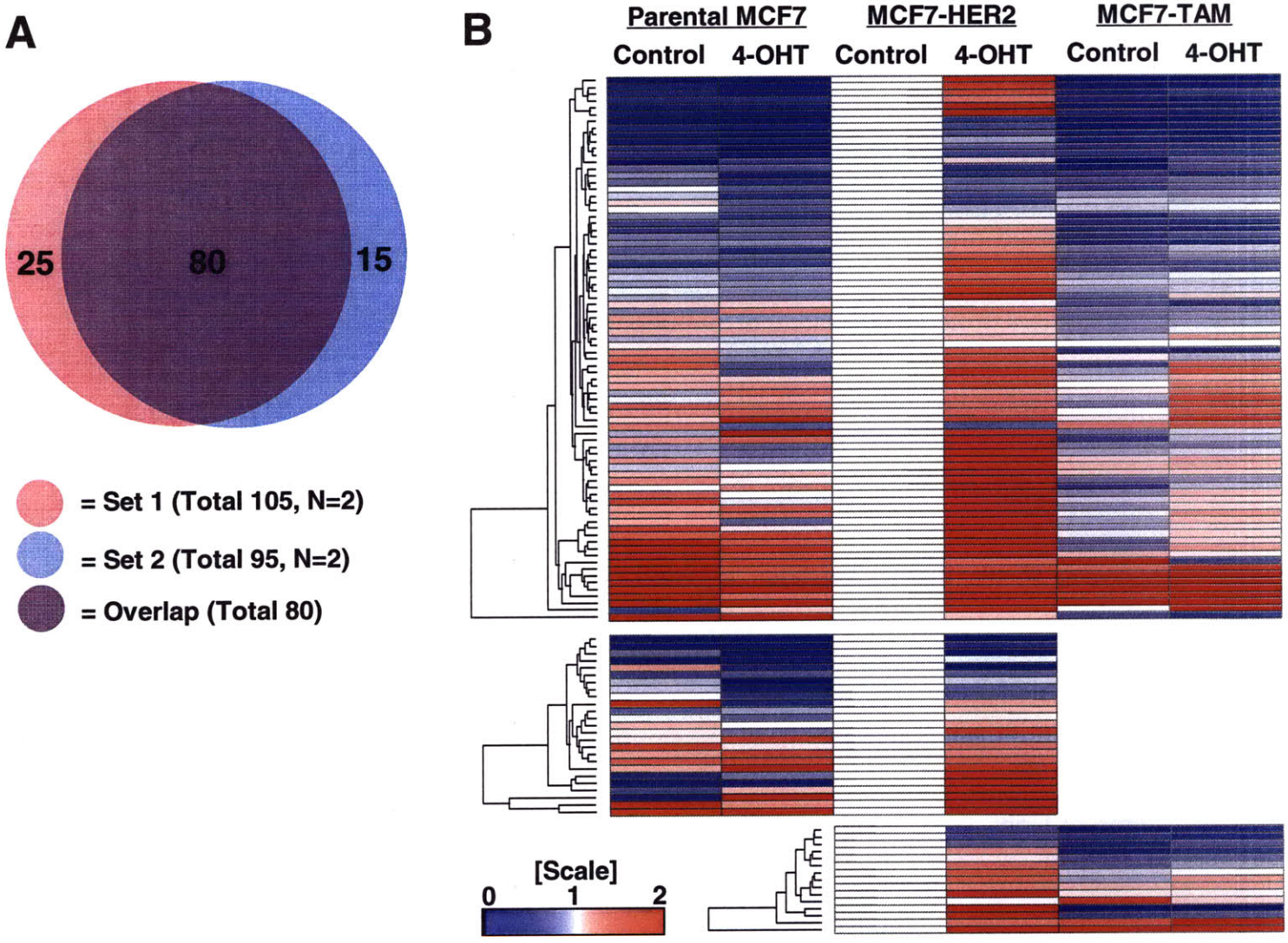


Figure 9. Results of MS-based quantitative proteomic analysis of tyrosine phosphorylation. (A) Of the 120 tyrosine phosphorylation sites identified and quantified in this analysis, 80 phosphorylation sites were quantified for all six conditions, while 25 were quantified in only the MCF7-parental and MCF7-HER2 cells (Set 1), and 15 were quantified in only the MCF7-TAM and MCF7-HER2 cells (Set 2). (B) Hierarchical clustering was performed to visualize the effects of 4-OHT treatment on tyrosine phosphorylation network across the three cell lines. Cell

context dependent changes are evident, as each of the cell lines displays a significantly different response to 4-OHT exposure. Heat map scale is fold-change relative to MCF7-HER2 control.

Data generated in this study indicate that exposure to 4-OHT affects tyrosine phosphorylation levels of many proteins and kinases and that this effect is strongly dependent on cellular context. For instance, as depicted in Figure 9B, exposure of MCF7-HER2 cells to 100 nM 4-OHT resulted in an increase in many tyrosine phosphorylation sites, while parental MCF7 cells displayed a much different response to the same treatment, with multiple phosphorylation sites decreasing in intensity while others displayed minimal change. MCF7-TAM cells are the most resistant to Tamoxifen and had the smallest response to 4-OHT exposure, with most sites demonstrating no change in phosphorylation, and only a select subset of phosphorylation sites increasing in intensity with 4-OHT.

The complex, context-dependent response to 4-OHT treatment was detected at both the network level and for individual proteins. To better visualize this cell-specific differential response to 4-OHT, quantification of protein phosphorylation changes in the canonical ErbB signaling network and downstream cell adhesion

proteins are depicted in Figure 10-12.

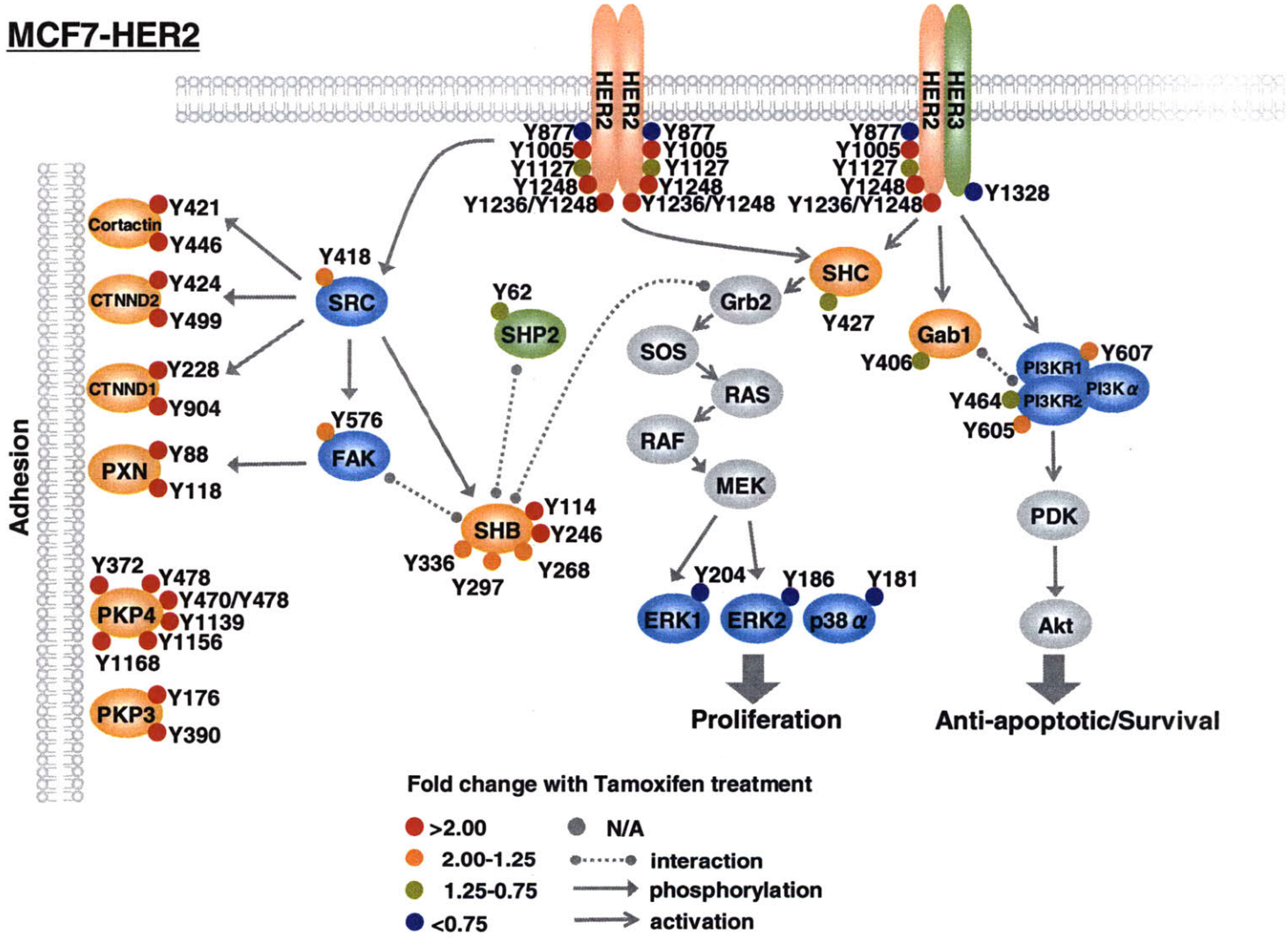


Figure 10. Schematic representation of the fold change in phosphorylation level in the canonical ErbB and cell adhesion signaling networks in response to 4-OHT treatment in MCF7-HER2. Note that many of the phosphorylation events increase significantly in response to 4-OHT exposure. For each phosphorylation site, the fold change in phosphorylation level relative to control treatment for each cell line is represented by red (>2-fold change), orange (1.25-2.00 fold change), green (0.75-1.25 fold change), or blue (<0.5-fold change).

As can be seen in these panels, five tyrosine phosphorylation sites were identified and quantified on the HER2 receptor tyrosine kinase. Of these five sites, three (pY1005, pY1236, and pY1248) demonstrated an increase in phosphorylation level, one (pY1127) was unchanged, and one (pY877) decreased in phosphorylation level following exposure of the MCF7-HER2 cells to 4-OHT (Figure 10 and Table 1). For HER2 pY1248, the C-terminal docking sites for many adaptor molecules, the basal phosphorylation level was ~30 fold higher in MCF7-HER2 cells compared to parental MCF7 and MCF7-TAM; in response to Tamoxifen treatment, the phosphorylation level increased by 2-fold in the MCF7-HER2 cells, as previously reported (Shou et al., 2004), but remained unchanged or slightly decreased in the parental MCF7 and MCF7-TAM cells (Figure 11-12, and Table 1).

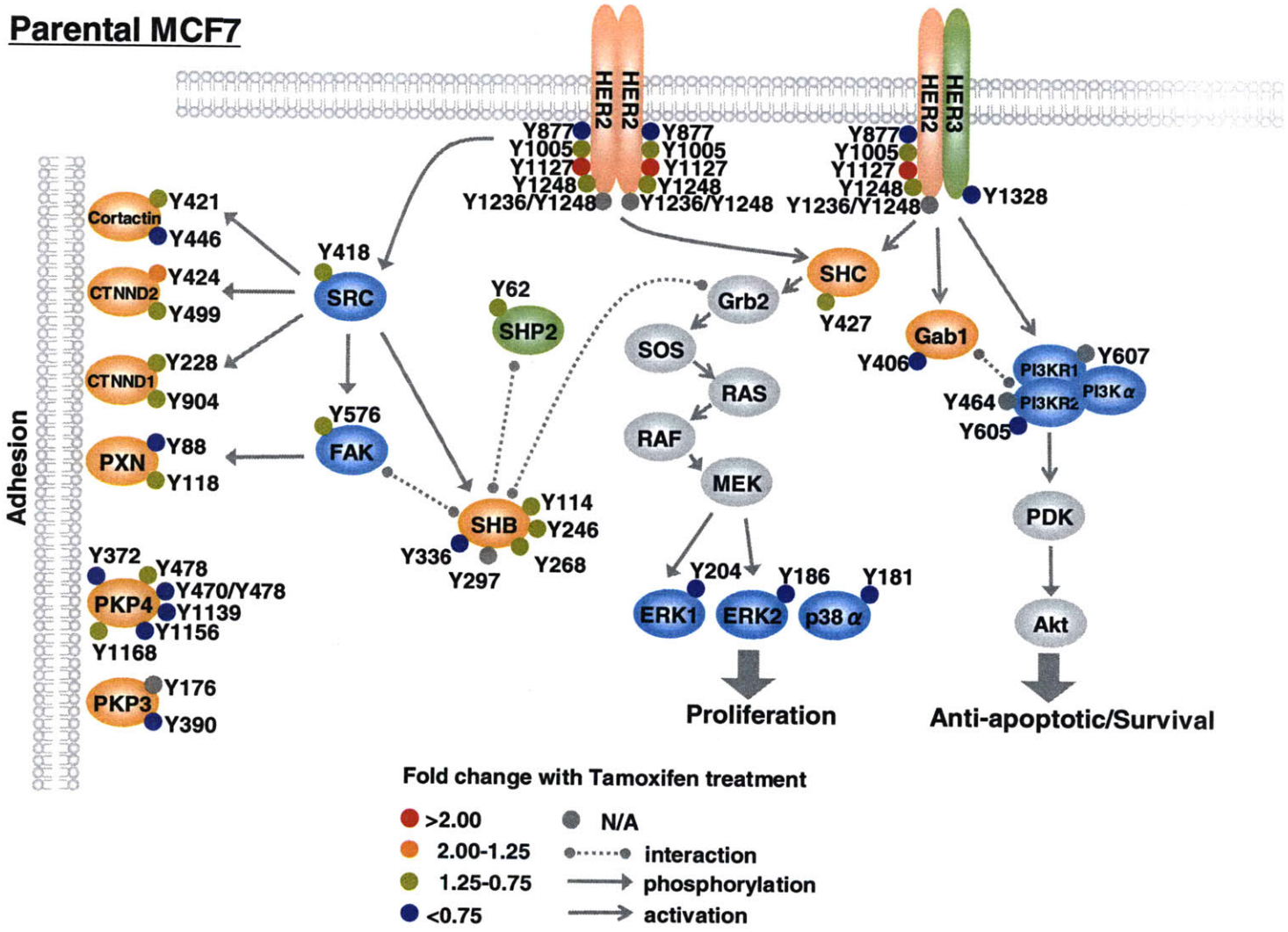


Figure 11. Schematic representation of the fold change in phosphorylation level in the canonical ErbB and cell adhesion signaling networks in response to 4-OHT treatment in parental MCF7. Note that many of the phosphorylation events increase significantly in response to 4-OHT exposure in MCF7-HER2, while most of the sites decrease in phosphorylation following exposure to 4-OHT in the parental MCF7 cells. For each phosphorylation site, the fold change in phosphorylation level relative to control treatment for each cell line is represented by red (>2-fold change), orange (1.25-2.00 fold change), green (0.75-1.25 fold change), or blue (<0.5-fold change).

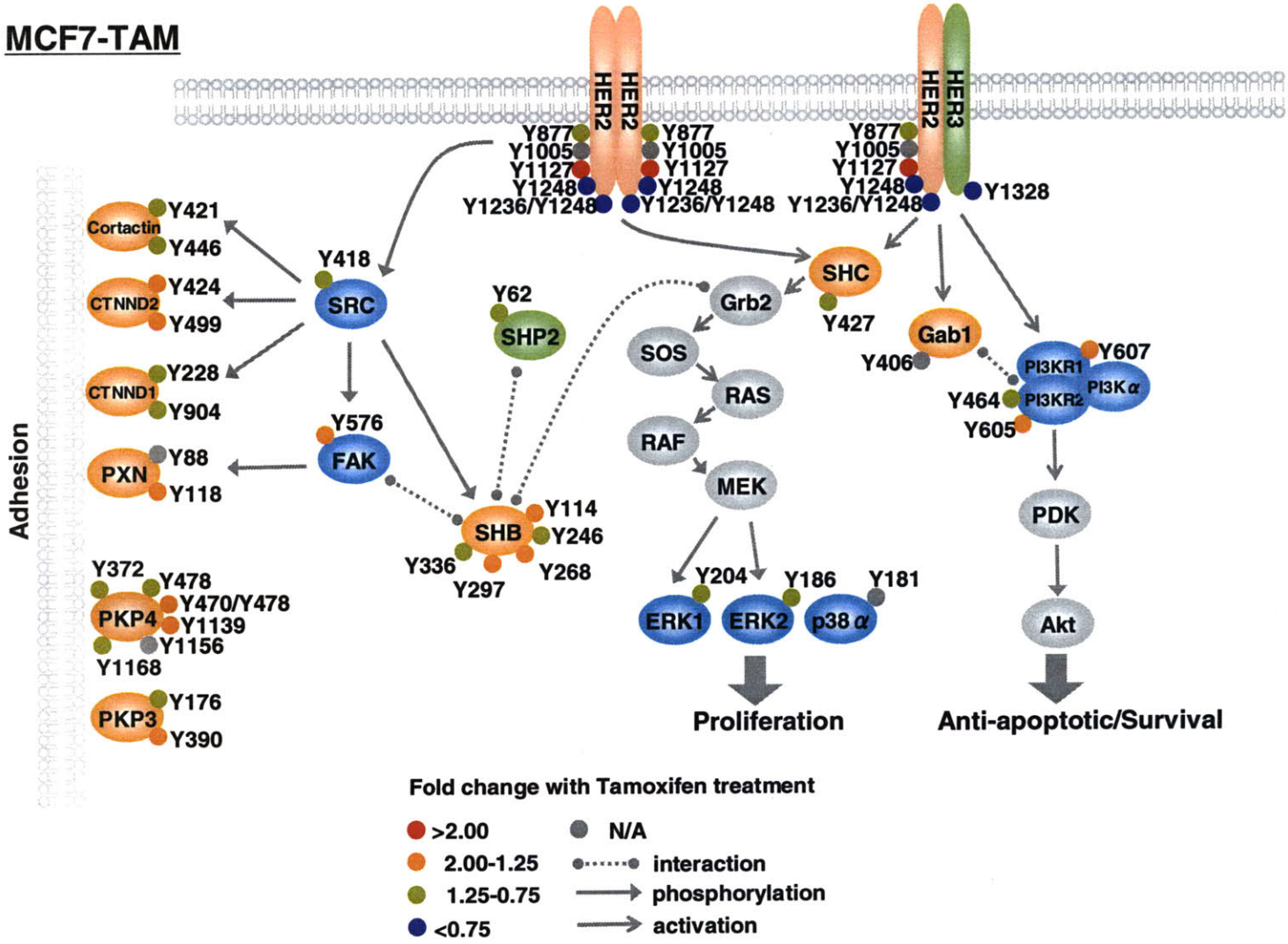


Figure 12. Schematic representation of the fold change in phosphorylation level in the canonical ErbB and cell adhesion signaling networks in response to 4-OHT treatment in MCF7-TAM. The MCF7-TAM cells display a more muted response, with most sites being unaffected by 4-OHT exposure, and selected sites increasing between 1.25 and 2-fold relative to the untreated MCF7-TAM cells. For each phosphorylation site, the fold change in phosphorylation level relative to control treatment for each cell line is represented by red (>2-fold change), orange (1.25-2.00 fold change), green (0.75-1.25 fold change), or blue (<0.5-fold change).

Phosphorylation levels for this site were also measured by ELISA, confirming the observed increase in HER2 pY1248 in MCF7-HER2 after 4-OHT treatment (Figure 13A). The relative increase in phosphorylation was not due to a change in HER2 expression level, as the total HER2 level was not significantly altered by exposure to 4-OHT (Figure 13B).

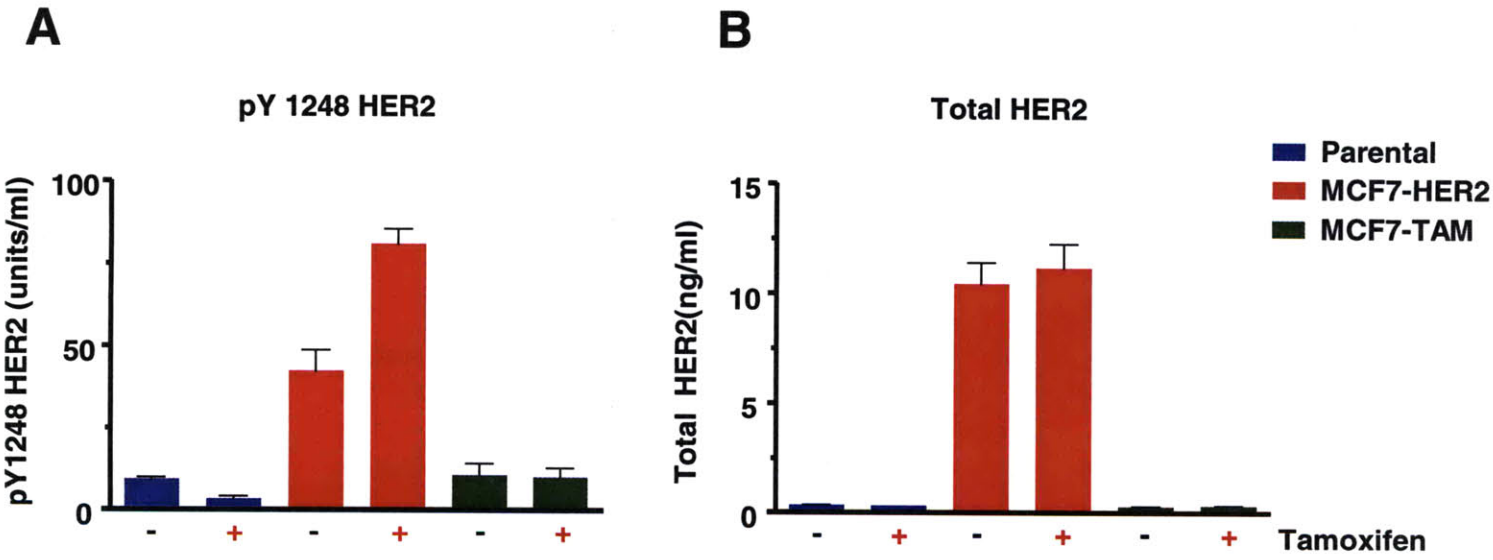


Figure 13. Enzyme linked immunosorbant assay (ELISA) measurement of HER2. (A) HER2 phosphotyrosine 1248 was quantified by ELISA in the three cell lines in the presence and absence of 4-OHT. Quantification pattern on HER2 pY1248 from ELISA is similar to the relative quantification results from MS-based phosphotyrosine level measurements. (B) As demonstrated by western blots in Supplementary Figure 1A, ELISA quantification also shows that the total HER2 expression level is similar for parental MCF7 and MCF7-TAM cells, but is increased ~40x in the MCF7-HER2 cells.

The cell context-specific 4-OHT response is further highlighted by HER2 pY1127, a site of unknown function, whose phosphorylation level increased dramatically in the MCF7 parental and MCF7-TAM cell lines in response to 4-OHT exposure, but was unchanged in the MCF7-HER2 cells. As figures clustering and network figure indicate, similar signaling complexity is recapitulated throughout the network. In order to understand how the complex regulation of these signaling networks relates to Tamoxifen resistance, it is necessary to identify those phosphorylation events that correlate most strongly with the primary phenotype of interest, growth rate in the presence of 4-OHT.

II.3.5 Human primary tumor analysis.

In addition to Tamoxifen resistance cell line models, we extended phosphoproteomic analyses of tyrosine phosphorylation events to primary human tumors. We received two primary human breast tumor samples from Dr. Morag Park at McGill University in Canada. These two tumors are selected from the McGill Univeristy Hospital tissue bank, the largest tumor data banks in Canada.

The first sample was from a patient who was diagnosed with ER^+ invasive ductal carcinoma (IDC) in 1991. The standard therapeutic procedure for such patients in 1991 was to surgically remove breast tumors, then put patients on five years of adjuvant Tamoxifen treatment. The first patient received the five years Tamoxifen treatment until 1996, and adjuvant Tamoxifen treatments were completed at this point. Unfortunately, this patient relapsed with grade 3^+ recurrent breast tumor with ER^+ and $HER2^+$ in 2000. The tumor sample was collected at the time of recurrence in 2000, when patient was at age 72. The second primary tumor sample was from a patient who was diagnosed as IDC with ER^+ and $HER2^+$ at grade 3^+ and tumor sample was collected at the time of diagnosis in 1999, when the patient was age at 77, before the start of adjuvant Tamoxifen treatments.

It would have been ideal to have the patient-matched samples. For example, the first sample at the time of diagnosis in 1991 would be perfect to compare changes between original and recurrent tumor samples. Unfortunately, such patient-matched samples were not available from the tumor bank, and the second tumor sample described above was chosen as the best histological match to the first primary tumor. To a limited extent, these tumor pairs are analogous to

the MCF7-HER2 cell lines with or without Tamoxifen treatment for 15 days because both data sets are for $ER^+/HER2^+$ samples with or without Tamoxifen treatment.

Both tumor samples were homogenized and processed to tryptic peptides in the same way as cell lysates were processed for MS analyses. Detection of tyrosine phosphorylation sites from the tumor xenografts has been difficult due to low levels of phosphotyrosine (Paul Huang, personal communication), and similarly, detection of phosphotyrosine from primary tumor samples was more difficult than detection from cell line samples. This is primary due to the heterogeneity of human tissues, in which stromal cells and fibroblasts are present in addition to cancer cells. In order to maximize the level of phosphotyrosine injected into mass spectrometer, the amount of peptides that are labeled with iTRAQ reagents were doubled, and analytical replicates were conducted within one MS analysis. Peptides from the recurrent tumor sample were labeled with iTRAQ 114 and 116, whereas peptides from the before-treatment sample were labeled with iTRAQ 115 and 117, enabling the calculation of average and standard error for all of the phosphorylation sites from a single MS analysis.

Forty tyrosine phosphorylation sites were identified and quantified, enabling

comparisons between a recurrence sample and a before-treatment sample (Figure 14). These 40 phosphorylation sites include sites from many signaling molecules, kinases, and phosphatases such as STAT2, PKCδ, PI3K regulatory subunit 1, WASP, DYRK1, FRK, FYN, HER2, HER3, PTPRA, FAK, SHIP2, MAPKs, and SHC. Intriguingly, the levels of phosphorylation from the majority of phosphorylation sites were higher in the recurrent tumor sample, and half of phosphorylation sites were more than two fold higher in the recurrent tumor sample. Importantly, phosphorylation sites on many tyrosine kinases were over two fold higher in the recurrent tumor sample including SRC family kinases, HER2, HER3, and FAK. In addition to tyrosine kinases, four MAP Kinases (ERK1, ERK2, p38α, and p38δ) were phosphorylated between four- and six-fold higher in a recurrent sample. Trends in the phosphorylation of SRC/FAK and ERK 1/2 were similar to those observed in phosphoproteomic analyses from MCF7-HER2 cell line models with Tamoxifen treatment. SRC/FAK and ERK1/2 pathways may have higher activities in a recurrent tumor compared to a before-treatment tumor in a similar manner that Tamoxifen treated MCF7-HER2 may have higher SRC/FAK and ERK1/2 pathways than untreated MCF7-HER2.

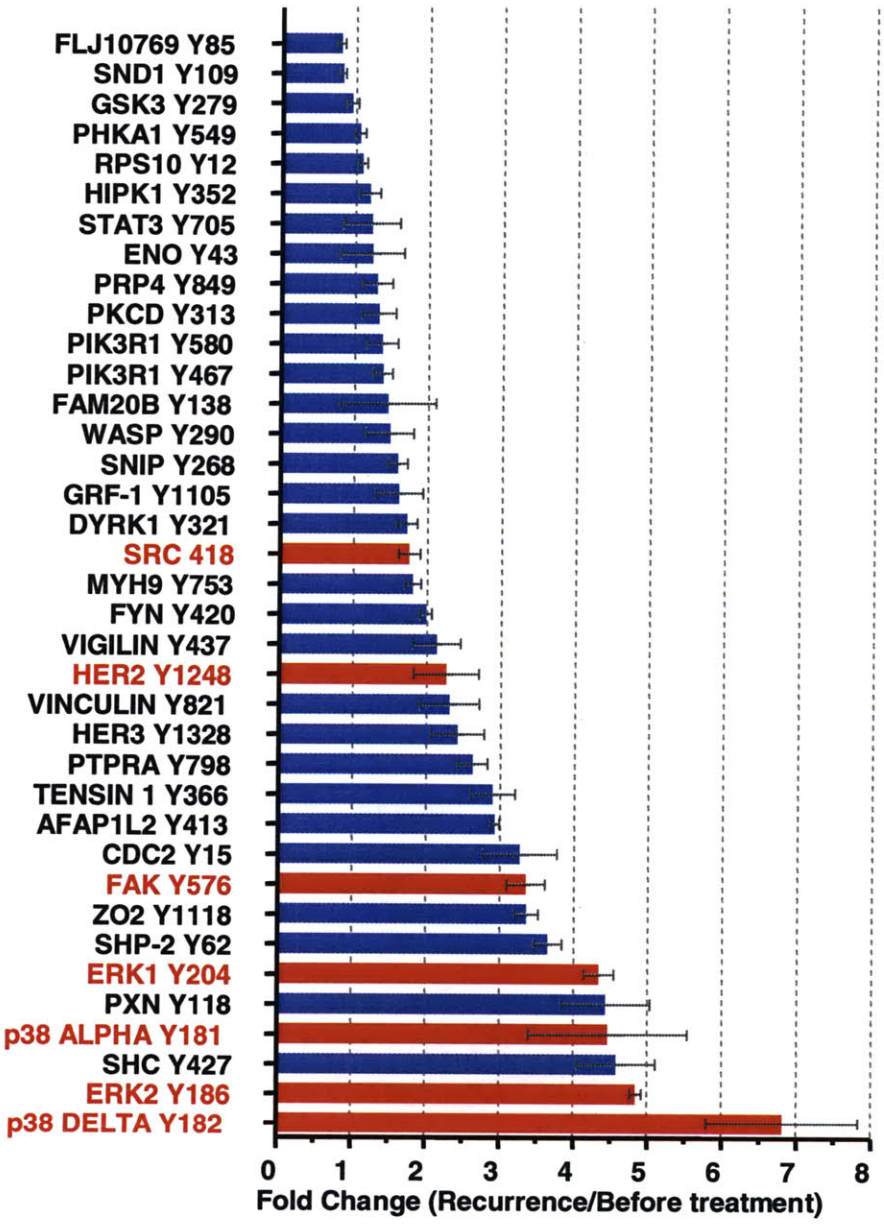


Figure 14. Phosphoproteomic analysis of human primary breast tumor. Forty tyrosine phosphorylation sites are identified. Fold changes between two primary tumor samples are shown (a sample from the patient with recurrent tumor over a sample from patient before Tamoxifen treatment. Note that HER2, Src, and FAK tyrosine kinases and four MAP Kinases (ERK1, ERK2, p38α, p38δ) were highly phosphorylated in a recurrent sample in a consistent manner as in results from MCF7-HER2.

II3.6 PI3K p85β R2 pY605 correlates with Akt pS473 levels and is highly phosphorylated in Tamoxifen resistant cells.

To identify those sites correlated with cell growth rate in the presence and absence of 4-OHT, we first checked phosphorylation levels on proteins known to be associated with cell survival/anti-apoptosis signaling. Notably, pY605 on phosphoinositide-3 kinase p85β regulatory subunit 2 (PI3K p85β R2) exhibited a 60% decrease in phosphorylation in parental MCF7, whereas the same site increased by ~70% and ~100% in MCF7-HER2 and MCF7-TAM, respectively, in response to 100 nM 4-OHT exposure (Figure 15). Although the function of the pY605 site is not known, a motif characterized by the surrounding amino acid sequence matches closely to the motif established for binding to the SH2 domain of the PI3K p85β regulatory subunit 1 (Scansite score=0.142) (Obenauer et al., 2003). Interaction between the two PI3K p85β regulatory subunits is a prerequisite for recruitment of PI3K p110α catalytic subunit which fully activates the trimetric class I PI3K (Neri et al., 2002). PI3K activation increases phosphatidylinositol 3,4

bisphosphate (PtdIns(3,4)P2) and phosphatidylinositol 3,4,5-triphosphate (PtdIns(3,4,5)P3) in the cellular membrane leading to recruitment and activation of Akt/PKB, which results in propagation of pro-survival signals. Based on the increased phosphorylation of PI3K p85β R2 pY605, and the potential role for this site in mediating the interaction with PI3K p85B R1, we hypothesized that the PI3K-Akt pathway may be differentially regulated between Tamoxifen sensitive (parental MCF7) and resistant (MCF7-HER2 and MCF7-TAM) cells, with Akt activation providing a key pro-survival signal in the Tamoxifen resistant cell lines treated with 4-OHT. To test the correlation between PI3K p85β R2 pY605 and Akt activation, Akt pS473, a surrogate for Akt activity, and total Akt were quantified by ELISA (Figure 15B). While total Akt expression was essentially unchanged across the three cell lines, Akt pS473 increased strongly in MCF7-HER2 and MCF7-TAM, and decreased slightly in parental MCF7 with 4-OHT treatment, correlating in trend, if not magnitude, with PI3K p85β R2 pY605. It is likely that the difference in magnitude between these signals may be due to other components of this signaling module, including PTEN, the lipid phosphatase regulating accumulation of PtdIns(3,4,5)P3. We have recently found that PTEN S-nitrosation levels are

increased in MCF7-TAM cells following 4-OHT exposure (data not shown), potentially decreasing the activity of this phosphatase and thereby amplifying the pro-survival PI3K-mediated Akt activation in the MCF7-TAM cells treated with 4-OHT.

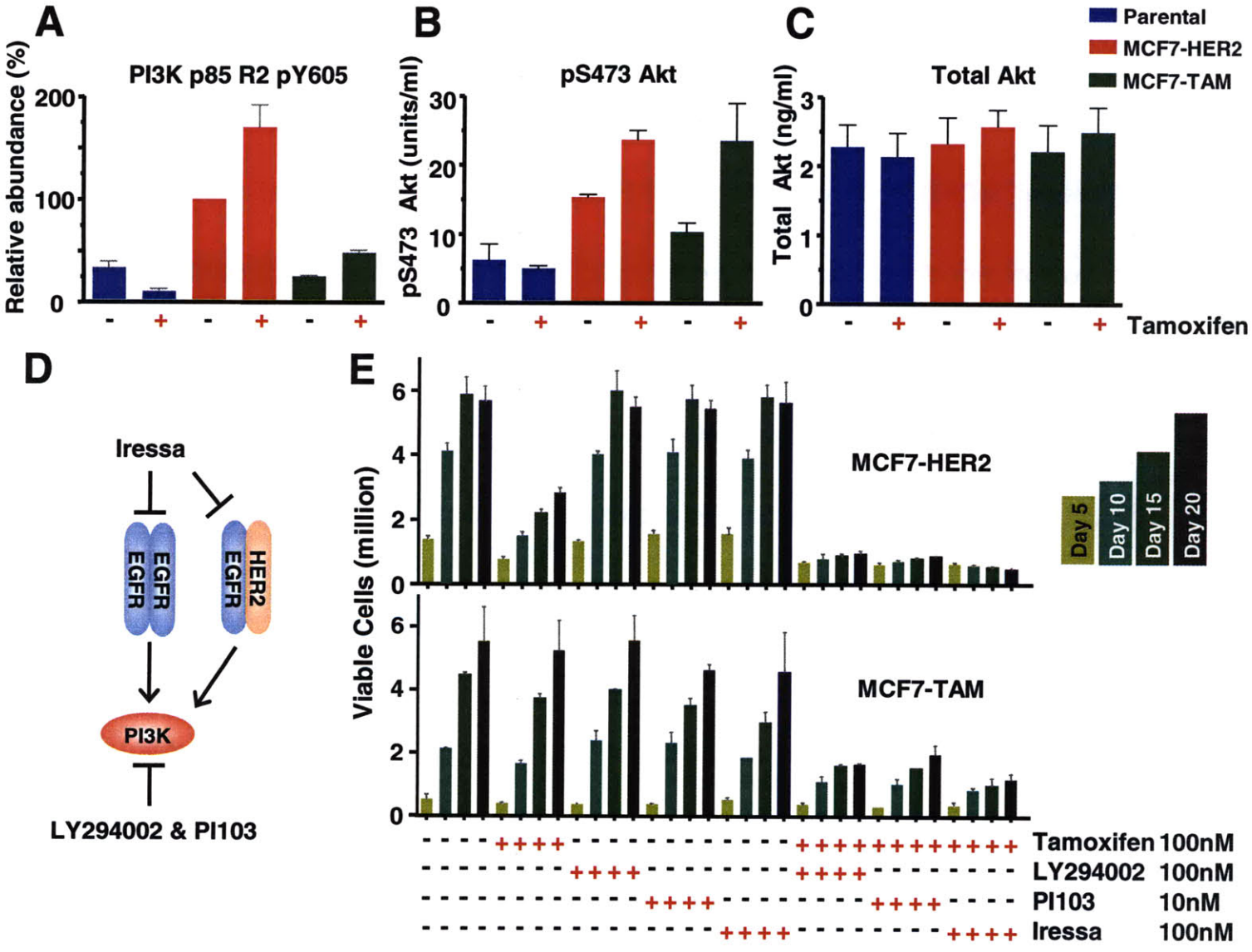


Figure 15. PI3K-Akt pathway activation and inhibition results. (A) Phosphorylation of pY605 on phosphoinositide-3 kinase p85β regulatory subunit 2 (PI3K p85β R2) increases in response to 4-OHT exposure in the Tamoxifen resistant MCF7-TAM and MCF7-HER2 cells, but decreases in the Tamoxifen sensitive MCF7-parental cells. (B) PI3K p85β R2 pY605 phosphorylation correlates with Akt phosphoserine 473, as measured by ELISA. (C) Change in phosphorylation is not associated with change in protein expression level, as quantified by ELISA measurement of total Akt expression level. (D) Schematic representation of the small molecule inhibitor strategy for PI3K inhibition. (E) Individual or combinatory treatment of Tamoxifen resistant cell lines with Tamoxifen and small molecules targeting PI3K pathway. Note that either of the PI3K inhibitors, when combined with 4-OHT, has a much greater effect on the MCF7-HER2 cells relative to the MCF7-TAM cells.

II.3.7 PI3K inhibition in the presence of Tamoxifen decreases growth rate of Tamoxifen resistant cells.

To determine the functional consequence of increased PI3K/Akt pathway activation in the Tamoxifen resistant MCF7-HER2 and MCF7-TAM cells treated with 4-OHT, small-molecule inhibitors targeting PI3K (LY294002, a PI3K inhibitor known to target all classes of PI3K, and PI103, a PI3K/mTOR dual-specificity inhibitor(Fan et al., 2006)) were used to modulate activity of this pathway in the presence and absence of 4-OHT (Figure 15D). As a positive control for this study, gefitinib

(Iressa, ZD1839), an EGFR kinase inhibitor that has been shown to revert Tamoxifen resistance in MCF7 cells overexpressing EGFR/HER2 (Shou et al., 2004) was used to modulate EGFR/HER2 activity, thereby decreasing PI3K activation indirectly. The concentration of each small-molecule inhibitor was titrated such that the inhibitor alone did not have an effect on cell growth rate in the absence of 4-OHT. As previously demonstrated, 4-OHT alone led to sustained cell growth for MCF7-TAM cells and sustained, but decreased, cell growth for MCF7-HER2. When cells were treated with the combination of 4-OHT and either PI3K inhibitor or gefitinib, a significant decrease in cell growth rate was seen for MCF7-HER2 and MCF7-TAM cells. Specifically, addition of LY294002 or PI103 with 4-OHT significantly reduced cell growth rate compared to 4-OHT alone, whereas addition of gefitinib with 4-OHT was more effective than either PI3K inhibitor, presumably due to the inhibition of other EGFR mediated pathways in addition to the PI3K-Akt pathway. Co-treatment with 4-OHT and PI3K inhibitors or gefitinib was more effective in the MCF7-HER2 cells compared to the MCF7-TAM cells, indicating the presence of a non-ErbB mediated Tamoxifen resistance mechanism in the MCF7-TAM cells. Although co-treatment with any of the three compounds

was moderately effective against MCF7-TAM, the effect of the co-treatment was not sufficient to cease cell growth, and therefore these conditions may facilitate selection of more resistant cells.

II.3.8 Erk1/2 phosphorylation pattern correlates with cell growth rate.

Since PI3K inhibition was insufficient to arrest cell growth rate of the MCF7-TAM cells exposed to 4-OHT, it was necessary to identify alternative targets in the signaling network that might be more effective at reverting Tamoxifen resistance of these cells. To this end, we observed that tyrosine phosphorylation on the kinase activation loop of the Erk 1/2 mitogen-activated protein kinases (MAPK) (Erk1 pY204 and Erk2 pY186) correlated with the phenotypic outcomes of the Tamoxifen sensitive and resistance MCF7 (Figure 16).

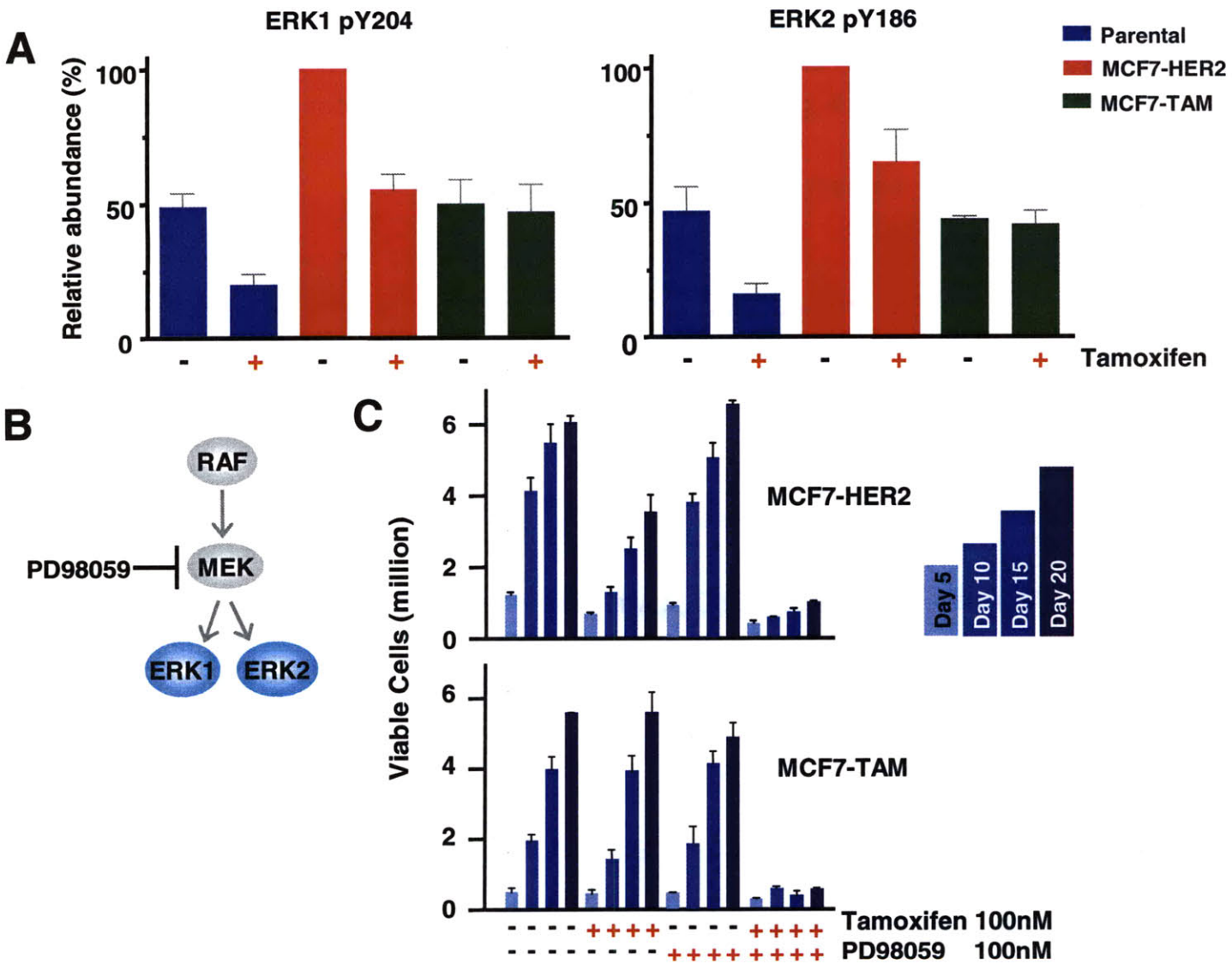


Figure 16. MEK-Erk1/2 pathway activation and inhibition results. (A) Phosphorylation of pY204 or pY186 on Erk1 or Erk2, respectively, correlates with cell growth rate for each of the cell lines with or without exposure to 4-OHT. (B) Schematic representation of the small molecule inhibitor strategy for MEK/Erk pathway inhibition. (C) Individual or combinatory treatment of tamoxifen resistant cell lines with tamoxifen and PD98059, a small molecule kinase inhibitor of MEK. Note that the combination of PD98059 and 4-OHT has a greater effect on the

MCF7-TAM cells compared to the MCF7-HER2 cells.

Specifically, parental MCF7 and MCF7-TAM cells show similar levels of Erk1 pY204 and Erk2 pY186 in the absence of 4-OHT and grow at approximately the same rate, while these levels are approximately 2-fold higher in the MCF7-HER2 cells that grow significantly faster compared to the MCF7 or MCF7-TAM cells. The correlation between phosphorylation and phenotype also applies following exposure to 100nM 4-OHT, as Erk1 pY204 and Erk2 pY186 decrease by 50% in parental MCF7 and MCF7-HER2 cells that are strongly affected by 4-OHT, but do not change significantly in MCF7-TAM cells, where the growth rate decreases only slightly in response to the same treatment. Quantification by ELISA to measure doubly phosphorylated Erk1/2 (pT202/pY204 and pT184/pY186) was in good agreement with the MS data (Figure 17A), and the change in signal was not due to alteration in the total protein expression level of these kinases, as determined by quantitative ELISA (Figure 17B).

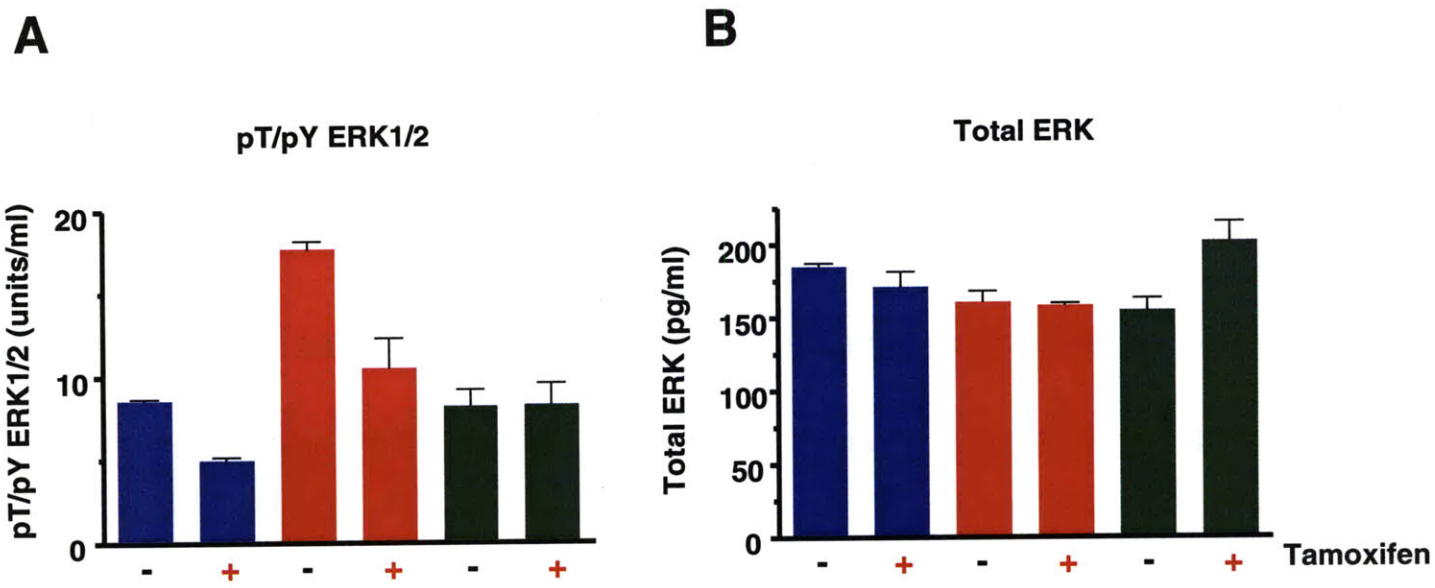


Figure 17. Enzyme linked immunosorbant assay (ELISA) measurement of ERK1/2. (A) Doubly phosphorylated ERK1 pT202/pY204 and ERK2 pT184/pY186 quantification pattern from ELISA is similar to the relative quantification results for the singly (tyrosine) phosphorylated peptides of ERK1 and ERK2 identified and quantified by MS-based phosphotyrosine measurements. (B) Total ERK1/2 expression level is relatively unchanged across the three cell lines and is unaffected by 4-OHT exposure.

To test the functional consequence of Erk1/2 activation on Tamoxifen resistance in these cells, we treated the cells with PD98059, a small molecule inhibitor of MAPK/Erk kinase (MEK), the kinase regulating Erk activation (schematically represented in Figure 16 B), and counted viable cells over a 20 day period following treatment with PD98059 or control in the presence and absence of 4-OHT (Figure

16C). As above, the concentration of PD98059 was adjusted such that minimal effect was seen when cells were treated in the absence of 4-OHT. Upon addition of 4-OHT in combination with PD98059, however, cell growth was greatly decreased in both the MCF7-HER2 and MCF7-TAM cells, albeit to different levels. Somewhat to our surprise, the cellular response to MEK/Erk pathway inhibition was different than that seen for PI3K/Akt pathway inhibition, as the MCF7-HER2 cells continued to grow at an appreciable rate upon MEK/Erk inhibition in the presence of 4-OHT, while cell growth was almost completely abrogated for the MCF7-TAM cells under these conditions (Figure 16C). The combined results of PI3K/Akt and MEK/Erk pathway inhibition indicate that the MCF7-HER2 cells may be more reliant on the PI3K pathway for survival and growth when exposed to 4-OHT, while the MCF7-TAM cells are much more reliant on the MEK/Erk pathway to maintain their Tamoxifen resistance. Neither inhibition strategy is sufficient to completely abrogate cell growth rate of both cell line resistance models.

II.3.9 Increase in Src and Src family kinase substrate phosphorylation with 4-OHT exposure

Although modulation of the PI3K and Erk pathways was sufficient to reduce Tamoxifen resistance in either the MCF7-HER2 or MCF7-TAM cells, the lack of a single treatment option that displayed high efficacy in both resistance mechanisms led us to further query the quantitative phosphorylation data to identify additional 4-OHT-induced phosphorylation changes that could represent alternate perturbation points to better modulate Tamoxifen resistance in both model cell lines. Perhaps the most striking effects of MCF7-HER2 exposure to 4-OHT are the phosphorylation changes induced on Src (note that the sequence of this phosphorylated peptide is identical for multiple Src family kinases, and therefore quantification of this site likely represents the protein expression-weighted average phosphorylation change on these Src family kinases expressed in these cells), Focal Adhesion Kinase (FAK), and cell adhesion proteins, including several known Src and FAK substrates (Figure 10 and Table 1). Although increased phosphorylation of Src in response to 4-OHT treatment has been previously reported in HER2-overexpressing Tamoxifen resistant MCF7 cells (Hiscox et al., 2006), here the unbiased network analysis reveals the functional consequence of

increased Src kinase activation through quantification of Src and FAK substrate phosphorylation. Among the known Src substrates, FAK is a tyrosine kinase known to localize at the focal adhesions and regulate cellular adhesion through phosphorylation of a variety of membrane proximal proteins (Hanks and Polte, 1997); moreover, FAK pY576 is a Src-dependent phosphorylation site and is indicative of FAK activity (Chaturvedi et al., 2007). Downstream of Src and FAK, multiple tyrosine phosphorylation sites on adhesion, cytoskeletal, and adaptor proteins increased phosphorylation in MCF7-HER2 cells exposed to 4-OHT (Figure 10). These hyperphosphorylated proteins include catenin δ1 (pY224 and pY902), catenin δ2 (pY424 and pY499), paxillin (pY118), plakophilin 3 (pY390), plakophilin 4 (pY372, pY470, pY478, pY1139 and pY1168), cortactin (pY421 and Y446), and SHB (pY114, pY246, pY268, pY297, pY336). Among these hyperphosphorylated proteins, catenin δ1, catenin δ2, and cortactin are previously described as Src kinase substrates (Mariner et al., 2001) (Huang et al., 1997). Paxillin Y118 has been characterized as a FAK phosphorylation site; phosphorylation of this site has been linked to the dissociation of paxillin from the focal adhesion protein complex, resulting in detachment of cells from adhesion sites (Bellis et al., 1995). SHB

phosphorylation also occurs in a Src-dependent manner and the SH2 domain of SHB is known to interact with phosphorylated FAK (Holmqvist et al., 2003). Although increased phosphorylation of the Src/FAK kinase and their substrates was absent in parental MCF7 exposed to 4-OHT (Figure 11), a moderate increase in phosphorylation of FAK and Src/FAK substrates was observed in MCF7-TAM upon 4-OHT exposure (Figure 12), indicating: (1) that increased Src/FAK activity may be associated with Tamoxifen resistance in MCF7, and (2) increased HER2 expression strongly potentiates Src/FAK kinase activation.

II.3.10 Inhibition of Src/Abl activity reverts Tamoxifen resistance

To test the role of Src family kinases in mediating Tamoxifen resistance, the Src/Abl dual-specificity inhibitor dasatinib (Lombardo et al., 2004) was used to decrease Src activity in the presence and absence of 4-OHT (Figure 18). The HSP90 inhibitors Geldenamycin and 17AAG were selected as a positive control for this experiment, as these compounds have been reported to be effective against Tamoxifen resistance in MCF7 cells (Beliakoff et al., 2003). In addition to their

direct effect on HSP90, these compounds should significantly alter HSP90 client proteins including HER2, ER, and Src (Whitesell and Lindquist, 2005), thereby perturbing the network at multiple levels simultaneously. As above, the concentration of each small-molecule inhibitor was titrated to establish the maximum concentration at which each inhibitor alone did not have an effect on cell viability (100nM for dasatinib and 10nM for geldenamycin and 17-AAG). Each inhibitor was then tested as a single agent or added in the presence of 100nM 4-OHT and cells were counted at multiple points over 20 days (Figure 18 B). While dasatinib, geldenamycin, or 17-AAG minimally affected cell growth when added to cells in the absence of Tamoxifen, addition of any of the three compounds in the presence of 100nM 4-OHT completely abrogated cell growth rate, such that the total number of cells was either unchanged or decreasing throughout the duration of the experiment. It is worth noting that the effect of Src/Abl inhibition by dasatinib was equally effective in reverting Tamoxifen resistance as the HSP90 inhibitors that were included as the positive control in this experiment. More importantly, dasatinib displayed good efficacy against both cell line models of Tamoxifen resistance, and may therefore represent a clinically relevant therapeutic option for reverting

Tamoxifen resistance regardless of HER2 expression status.

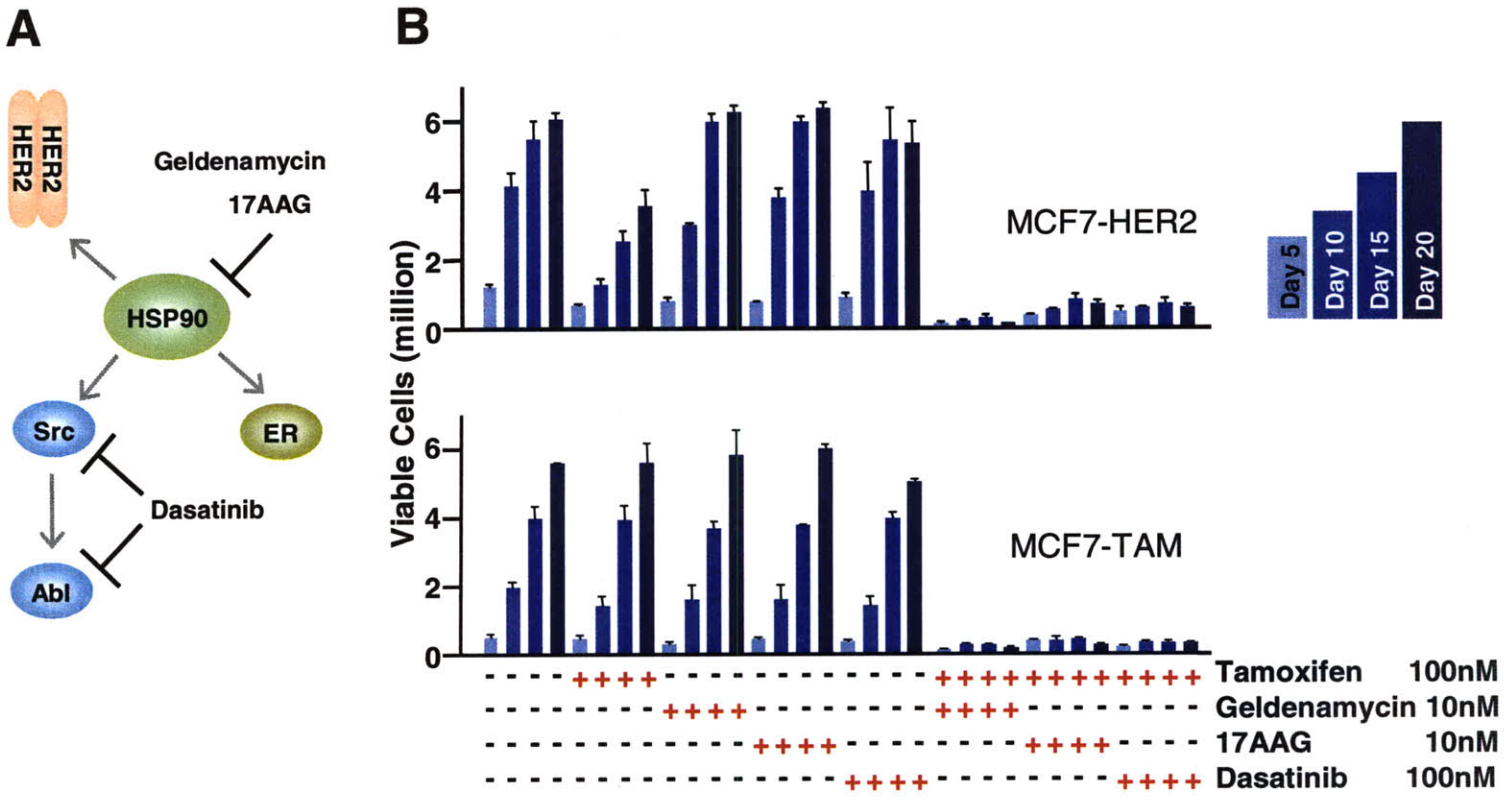


Figure 18. Reversion of tamoxifen resistance by HSP90 or Src/Abl inhibition. (A) Schematic representation of the targets of dasatinib, geldenamycin, and 17AAG. Note that HSP90 is a chaperone protein with multiple additional client proteins that are not represented in this schematic. (B) Combinatory treatment of tamoxifen resistant cell lines with tamoxifen and dasatinib, geldenamycin, or 17AAG leads to a significant decrease in the cell growth rate of the tamoxifen resistant cells, indicating reversion of tamoxifen resistance under these conditions. Note that Src/Abl inhibition works equally well against both cell lines, and has similar efficacy to the positive control HSP90 inhibitors (geldenamycin and 17AAG).

II.3.11 Analyses of tyrosine phosphorylation events in response to NOS

inhibitor treatments

One model for the basis of Tamoxifen resistance is downregulation of phosphatase activities, rather than kinase activation. In collaboration with the Tannenbaum lab, we analyzed the relationships between deregulated phosphatase activities and tyrosine phosphorylation events. Tamoxifen resistant cells express inducible nitric oxide synthase (iNOS) whereas Tamoxifen sensitive cells do not express iNOS (K. Pant, personal communication). Phosphatases such as PTEN and BDP-1 are S-nitrosated at cysteine residues in Tamoxifen resistant cells expressing iNOS. S-nitrosated phosphatases often become inactive due to the lack of functional, presumsbly catalytic, cysteine residues involved in their enzymatic reactions.

While 100nM Tamoxifen fails to suppress growth of Tamoxifen resistant cells, combinatorial treatment of Tamoxifen and NOS inhibitors affect growth of these cells (K. Pant, personal communication). Use of NOS inhibitors reversed the S-nitrosation on PTEN and BDP-1 in Tamoxifen resistant cells. As Tamoxifen treatment perturbs tyrosine phosphorylation events in Tamoxifen resistant cells, we

analyzed the effects of combinatorial treatment of Tamoxifen and a NOS inhibitor on tyrosine phosphorylation events. We employed two NOS inhibitors (1400W and S-methylthiocitrulline (SMTC)) at either 1mM or 0.3mM. Combinatorial treatments of 100nM Tamoxifen and 1mM 1400W or 0.3mM SMTC suppress growth of MCF7-TAM and MCF7-HER2 cells to a similar extent, while combinatorial treatment of 100nM Tamoxifen and 1mM SMTC has a stronger growth suppression effect (K. Pant, personal communication).

MCF7-TAM cells were treated with Tamoxifen alone, or Tamoxifen in combination with two different NOS inhibitors. MCF7-TAM cell samples were subjected to anti-pY immunoprecipitation-IMAC-LC-MS/MS analyses. Two analytical replicates were performed (Figure 19).

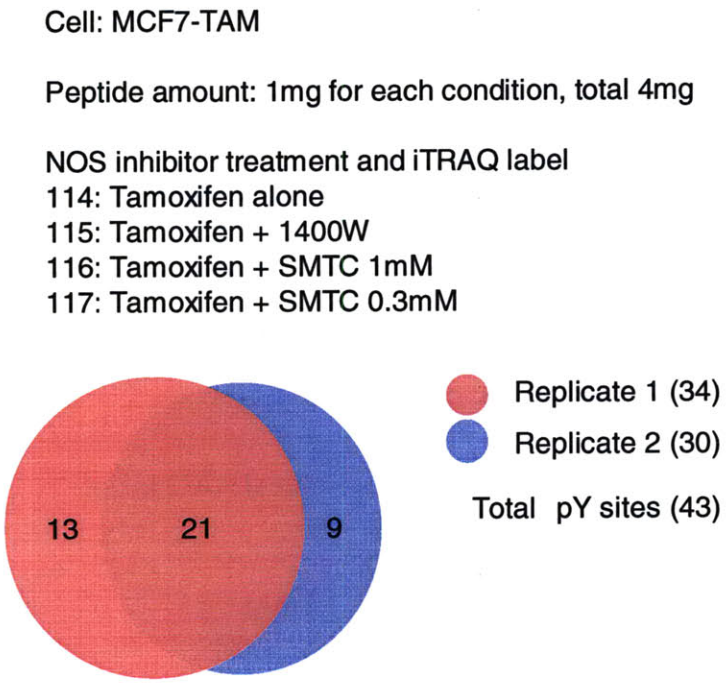


Figure 19. Summary of mass spec analyses for MCF7-TAM cells with Tamoxifen and NOS inhibitor treatments.

In total, 43 tyrosine phosphorylation sites were identified and quantified. Twenty-one phosphorylation sites were identified in both analyses, while 13 and 9 phosphorylation sites were identified in either analytical replicate. The number of tyrosine phosphorylation sites identified from each replicate is 34 and 30, indicating that basal tyrosine phosphorylation levels in MCF7-TAM are relatively low as typical mass spec analyses of cell lines yield higher numbers.

In addition to analyses of MCF7-TAM, tyrosine phosphorylation analyses of

MCF7-HER2 with combinatorial Tamoxifen and NOS inhibitor treatments were performed. MCF7-HER2 cell samples were subjected to anti-pY immunoprecipitation-IMAC-LC-MS/MS analyses. Two analytical replicates were performed (Figure 20).

Cell: MCF7-HER2

NOS inhibitor treatment and iTRAQ label
114: Tamoxifen alone
115: Tamoxifen + 1400W
116: Tamoxifen + SMTC 1mM
117: Tamoxifen + SMTC 0.3mM

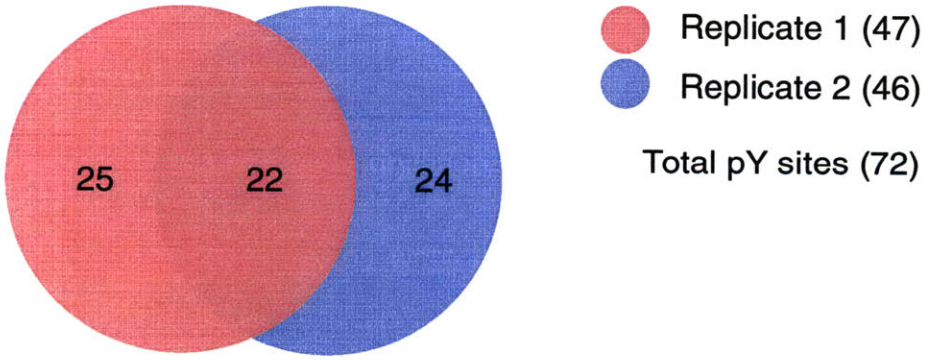


Figure 20. Summary of mass spec analyses for MCF7-HER2 cells with Tamoxifen and NOS inhibitor treatments.

A total of 72 tyrosine phosphorylation sites were identified and quantified. Twenty-two phosphorylation sites were identified in both analyses. Considering

approximately 60% reproducibility of mass spec analyses of identical samples in information dependent acquisition modes, 22 overlapping phosphorylation sites (47%) were lower than expected. 25 and 24 phosphorylation sites were identified in either analytical replicate. The number of tyrosine phosphorylation sites identified from each replicate is 47 and 46, indicating that basal tyrosine phosphorylation levels of MCF7-HER2 are higher relative to MCF7-TAM.

II.3.12 PI3K p85 pY605 increases with NOS inhibitor treatments in MCF7-TAM and MCF7-HER2.

The quantification of tyrosine phosphorylation analyses for MCF7-TAM and MCF7-HER2 treated with Tamoxifen and NOS inhibitors is listed in Table 2 and 3. The quantification were normalized to Tamoxifen treatement alone (iTRAQ114) in the all four analyses. Among the phosphorylation sites identified, PI3K p85 pY605 was detected in one of the two analytical replicates for MCF7-TAM data sets, and the phosphorylation site was detected in both analytical replicates for MCF7-HER2 data sets, enabling comparisons of the change at PI3K p85 pY605 in both

MCF7-TAM and MCF7-HER2. PI3K p85 pY605 was increased in all NOS inhibitor treatment conditions in MCF7-TAM cells (between 20-104%). Interestingly, PI3K p85 pY605 was also increased in MCF7-HER2 cells under all NOS inhibitor treatment conditions (between 67-79%). From the previous Scansite analyses, PI3K p85 pY605 is predicted to interact with the SH2 domain on another PI3K p85 subunit in order to recruit a PI3K catalytic subunit to fully activate the trimeric class I PI3K. Thus PI3K pY605 is a potential indicator for PI3K activity, with incrasing phosphorylation predicting increasing activity.

While further experimental validations are required, the possible consequences in the PI3K activity after Tamoxifen treatment compared to the combinatorial treatment could be described as below (Figure 21).

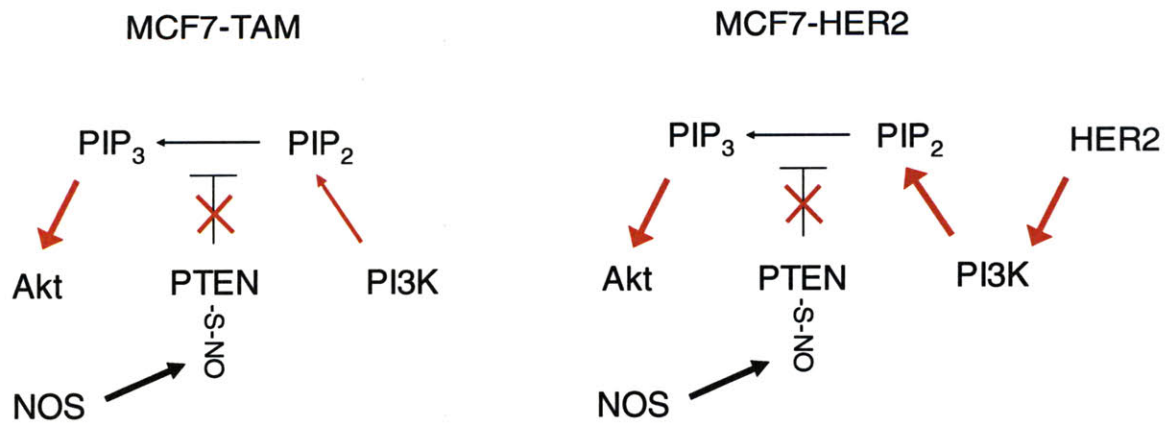


Figure 21. PI3K-Akt pathways in MCF7-TAM and MCF7-HER2 with Tamoxifen treatment.

HER2 expression strengthens PI3K activity in MCF7-HER2 cells. Tamoxifen treatment potentially increases PI3K activity in MCF7-TAM and MCF7-HER2, while such increase was not previously seen in the parental MCF7 cell line. Tamoxifen treatment induces iNOS expression and NOS production in MCF7-TAM and MCF7-HER2. Phosphatases including PTEN and BDP-1 are S-nitrosated at cysteine residues in response to iNOS expression and NOS production, resulting in potential decrease in phosphatase activities. Akt pS473 is increased in MCF7-TAM and MCF7-HER2 in response to Tamoxifen treatment as previously described in the ELISA experiments and western blots (K. Pant, personal communication). These data are consistent with the model show in figure 21.

Unlike Tamoxifen treatment alone, combinatorial treatments with Tamoxifen and NOS inhibitors may result in the following changes in PI3K pathways (Figure 22).

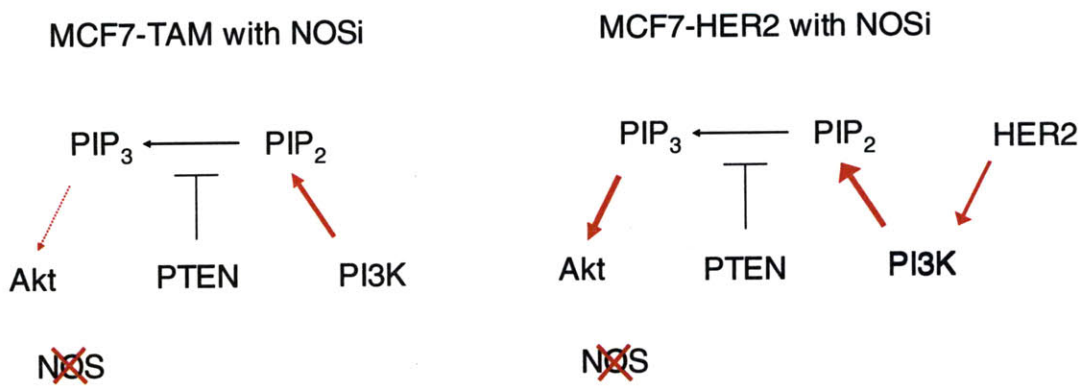


Figure 22. PI3K-Akt pathways in MCF7-TAM and MCF7-HER2 with Tamoxifen treatment and NOS inhibitor combinatorial treatments.

NO production decreases due to iNOS inhibition, resulting in potential recovery of phosphatase activities such as PTEN activity. Normal PTEN activity restores the balance between PIP_2 and PIP_3 levels, resulting in restored Akt activity in MCF7-TAM. In MCF7-HER2, restored PTEN activity may not be able to maintain the same balance between PIP_2 and PIP_3 levels because of potent signals from HER2 to activate a PI3K activity, resulting in minimal decrease in Akt phosphorylation. Akt activity measured by Akt pS473 decreases in response to NOS inhibitor treatment in MCF7-TAM, but not in the MCF7-HER2 (K. Pant, personal communication), indicating that NOS inhibition leads to restored Akt activity in MCF7-TAM but HER2 overexpression supersedes Akt regulation through PTEN S-nitrosation.

In summary, potential consequences of Tamoxifen and NOS inhibitor treatments on PI3K-Akt pathways are shown in Figure 23.

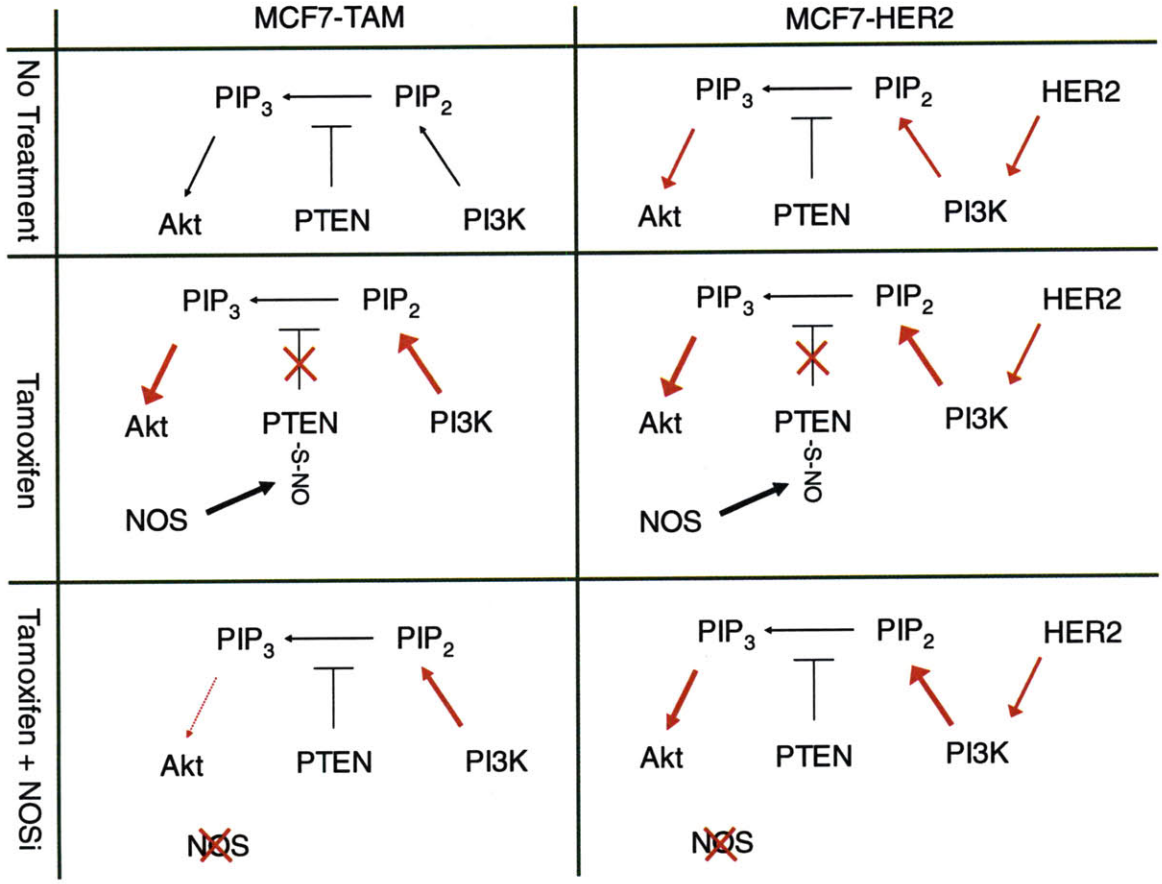


Figure 23. Summary of PI3K pathways in MCF7-TAM and MCF7-HER2 with Tamoxifen and NOS inhibitor treatments.

While multiple phosphorylation events observed in the data support the above model, additional biological replicates would greatly increase confidence in this model. Additionally, independent approaches to testing this model have been explored. Measurements of NO concentration inside Tamoxifen resistant cells

indicated that NO concentration was increased in response to Tamoxifen treatment (K.Pant, personal communication). In addition to NO measurements, identification of the S-nitrosated cysteine residues on phosphatases including PTEN and *in vitro* phosphatase assays provide evidence for the hypothesis that Tamoxifen treatment deactivates phosphatase activity via iNOS induction and NO productions.

II.3.13 Rescue of MCF7-HER2 cells from Tamoxifen treatment with growth factors

While MCF7-HER2 remains sensitive to Tamoxifen treatment and shows partial resistance to Tamoxifen compared to parental MCF7, it remained unclear what roles HER2 plays in addition to providing cells with growth signals and possibly cross-talking to ER pathways. Due to emerging reports of cross-talk between receptor tyrosine kinases that are not previously known to heterodimerize (Huang et al., 2007) (Rexer et al., 2009), I tested if receptor tyrosine kinases expressed in MCF7-HER2 cells show any sign of cross-talk. As MCF7 cells express high level of Insulin-like growth factor-1 receptor (IGF1-R), we conducted

experiments testing the effects of IGF1 treatments on parental MCF7 and MCF7-HER2, hypothesizing that HER2 overexpression may result in phenotypic differences when exposed to IGF1 treatments. We treated both parental MCF7 and MCF7-HER2 with 100nM or 1uM Tamoxifen, which lead to decreased viable cell numbers in both cells. In order to test if growth-promoting characteristics of IGF1 affect a number of viable cells, 100ng/ml (saturating dose) of IGF1 is added to the cell culture media (Figure 24). As additional conditions for the study, we used EGF and HRG as a stimulant of HER2 mediated signaling pathways at 100ng/ml (saturating dose). The aims of these experiments were to test if any of these growth factor treatments rescue either the parental MCF7 or MCF7-HER2 from Tamoxifen treatments.

Both 100nM and 1uM Tamoxifen could suppress the growth of parental MCF7. When EGF, HRG or IGF1 is added to media, parental cells showed a marginal increase in the number of the viable cell. As parental MCF7 express high levels of IGF1-R, low levels of EGFR and HER2, the results indicates that suppression of growth by Tamoxifen is more potent than the growth promoting effects from the IGF1, EGF or HRG treatments, although partial rescue was

observed.

In contrast, MCF7-HER2 responded to IGF1, EGF, and HRG and the number of the viable cells were the same as the control condition even with Tamoxifen treatment. The rescue effects are present in both 100nM and 1uM Tamoxifen treatment, indicating that a saturating dose of each growth factor treatments is more potent than the growth suppression from Tamoxifen treatments. We predicted that EGF and HRG treatments would have better rescue outcomes in MCF7-HER2 than parental MCF7 due to HER2 overexpression. However, we predicted that IGF1 treatment would have similar effects in MCF7-HER2 and parental MCF7 as IGF1-R expression are predicted to be similar in both cell types. Surprisingly, IGF1 treatment showed strong rescue effect in MCF7-HER2 cells but such rescue effect was not seen in parental MCF7.

We have not measured the expression levels of IGF1-R, EGFR, or HER3 in MCF7-HER2 cells relative to parental MCF7. As overexpression of a receptor tyrosine kinase is known to potentially affect expression levels of other receptor tyrosine kinases, the exact mechanisms through which IGF1 mediates rescue of MCF7-HER2 cells from Tamoxifen treatment remain unclear. It is possible that the

IGF1-R expression level is higher in MCF7-HER2 which caused potent IGF1-R signals. IGF1 is known to bind IGF1-R and Insulin receptor (IR) depending on expression levels of each receptor and doses of IGF1. The saturating dose of IGF1 used in this study potentially affects both IGF1-R and IR activation. The activated IGF1-R and IR signaling networks might be amplified by, and/or cross-talk to HER2 signaling network in the MCF7-HER2, leading to phenotypically different outcomes compared to parental MCF7. Another potentially relevant experiment is to conduct the same rescue experiments in the presence of insulin. The MCF7 cell line is known to grow better in the media containing insulin. In order to better understand the effects of each growth factor on MCF7-HER2, it may be important to conduct titration experiments. In addition, MS analyses of tyrosine phosphorylation events in MCF7-HER2 cells with growth factor stimulation may highlight potential cross-activations of receptor tyrosine kinase and downstream signaling events. As IGF1-R is an important therapeutic target in multiple malignancies, better understanding of potential cross-talk mechanisms between IGF1-R and HER2 are of significant interest.

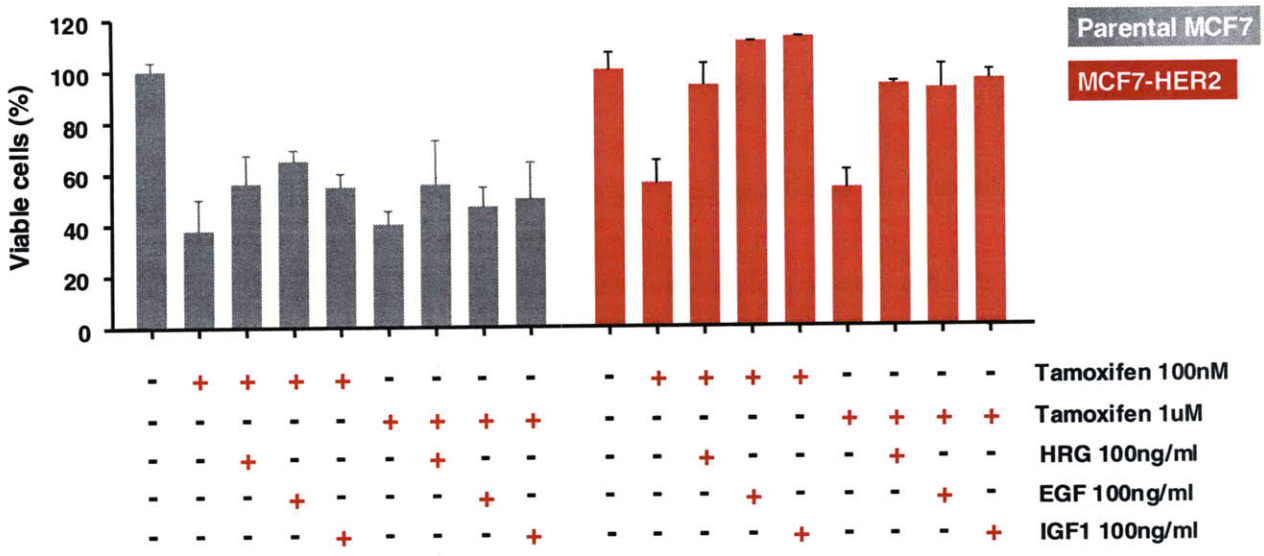
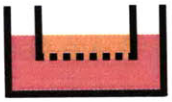


Figure 24. Rescue experiments with HRG, EGF1 and IGF1 on parental MCF7 and MCF7-HER2. Cells are counted at the day 9 for the parental MCF7 cells, and the day 6 for the MCF7-HER2 cells.

II.3.14 Migration and invasion abilities of Tamoxifen resistant cells

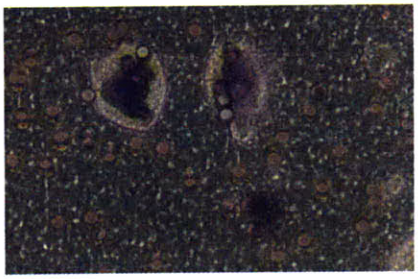
Adhesion molecules controlling cell-cell attachments and cellular locomotion are highly regulated by post-translational modification such as phosphorylation. Potential consequences of hyperphosphorylation on adhesion molecules in Tamoxifen resistant cell lines include increased migratory and invasive abilities. In order to test if Tamoxifen treatment makes Tamoxifen resistant cells more migratory and invasive, we have conducted transmembrane migration assays

and matrigel invasion assays with Boyden chambers. Both migration and invasion assays were conducted 24, 48, and 72 hours after seeding cells in Boyden chambers. There were no differences between time points, indicating that 24 hours is long enough for any mobile cells to pass through the holes in the membrane (data not shown). The lower chambers of each well are filled with the full growth media which contains chemoattractants present in FBS. Experimental design, a representative image and results from migration assays are shown in Figure 25.



Orange = 100ul Serum free DMEM with cells ($1x10^5$ cells/cm^2) (For invasion assay, Matrigel is 100ul of 1:30 diluted in serum free DMEM)

Pink = 500ul 10% FBS in DMEM



	Well 1	Well 2	Well 3	Ave.
MCF7-Parental, day 9, control	12	20	9	13.6
MCF7-HER2, day 6, control	22	28	31	27.0
MCF7-TAM, day 9, control	26	14	18	19.3
MCF7-Parental, day 15, Tamoxifen	21	20	13	18.0
MCF7-HER2, day 15, Tamoxifen	17	40	20	25.6
MCF7-TAM, day 15, Tamoxifen	21	11	16	19.3

Figure 25. Migration assay design and results. Experimental design of Boyden chamber, image of migrated cells at the bottom of membrane. The numbers are sum of measurements from two non-overlapping areas in each well.

To our surprise, the number of cells on the bottom side of the membrane was considerably low compared to the numbers expected similar membrane migration experiments with other breast cancer cell lines. Because the number of migrated cells was low, it was difficult to assess if Tamoxifen treatment made Tamoxifen resistant cells more migratory than vehicle treatment did. As the number of the migrated cells was low, the results from the invasion assay were even lower as cells not only have to pass membrane but also a matrigel layer (data not shown).

In order to add positive control conditions and test the validity of migration assay conditions, T47D and Met2A cell lines were tested for their migration ability in the same migration experimental settings (Figure 26). Both Met2A and T47D migrated through membrane, indicating that the protocol is valid for these two cell lines to show migration ability.

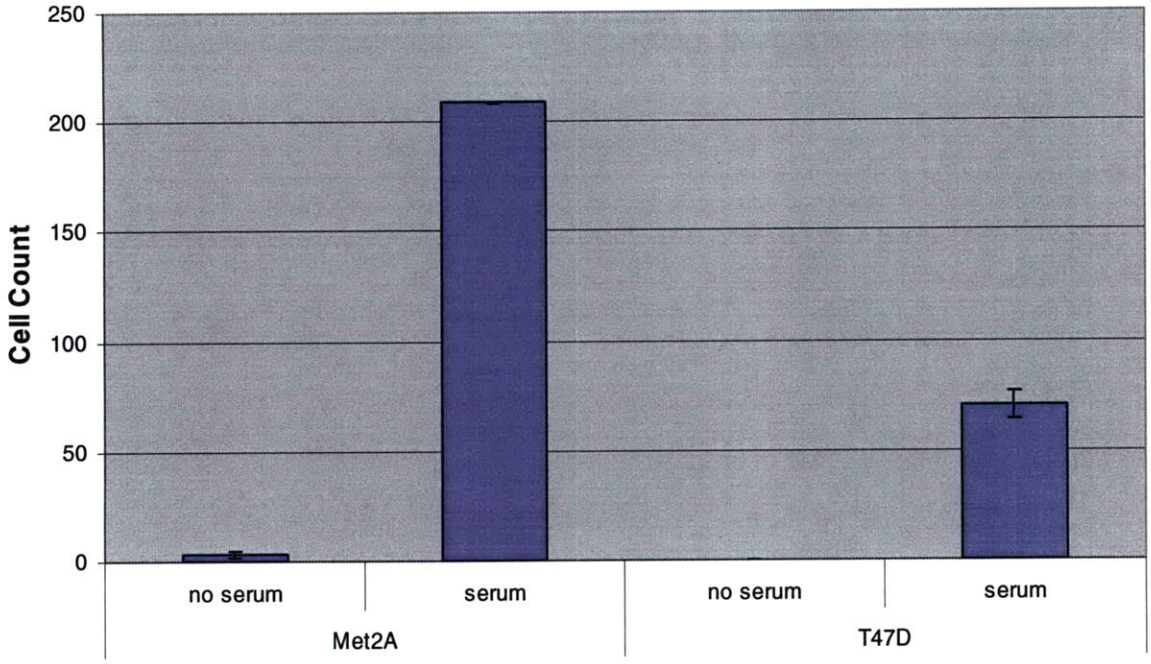


Figure 26. Migration assay with Met2A and T47D. Presence of 10% serum in bottom chamber makes both cell lines migrate. Note the higher number of migrated cells compared to MCF7 cells. Error bars are calculated from two wells in analytical replicates. Met2A with serum has an error bar, but too small to be visible in this figure.

MCF7 cells express high levels of E-cadherin, making them tightly anchored to each other. In order to re-test the migratory and invasive ability of Tamoxifen resistant cells in relatively favorable conditions for MCF7 cells to migrate, we used E-cadherin neutralizing antibodies to dissociate tightly anchored MCF7 cells during seeding of the top layer of the membrane during the migration assay. (Figure 27).

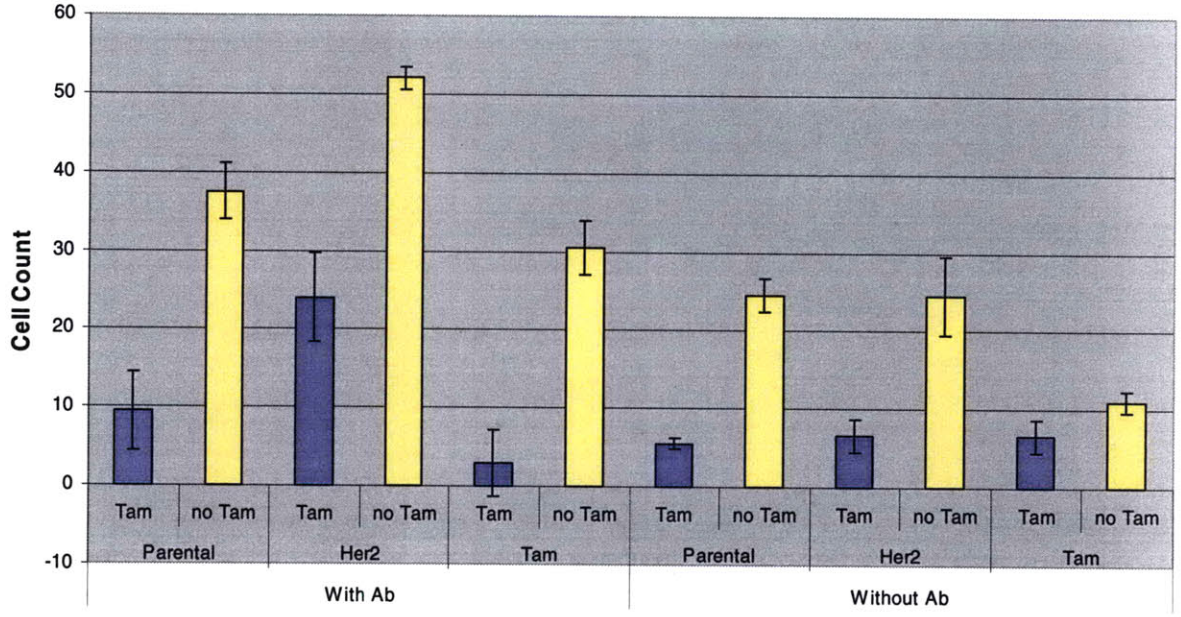


Figure 27. Migration assay with E-cadherin antibody treatment.

According to manufactures protocol for E-cadherin, 12ug/ml antibody would be sufficient for inhibitory experiments in cell culture. However, such treatment did not result in increase in migration of MCF7 cells (data not shown). We repeated experiments with 100ug/ml antibody, and obtained results above. E-cadherin treated MCF7 cells are more migratory than MCF7 cells without E-cadherin antibody treatments. Contrary to our expectation, Tamoxifen treated MCF7 cells were less migratory than control treated cells.

At 100ug/ml E-cadherin antibody treatments, there was a marginal increase

in the number of invasive MCF7 cells, with MCF7-HER2 cells being more invasive than parental MCF7 and MCF7-TAM (Figure 28).

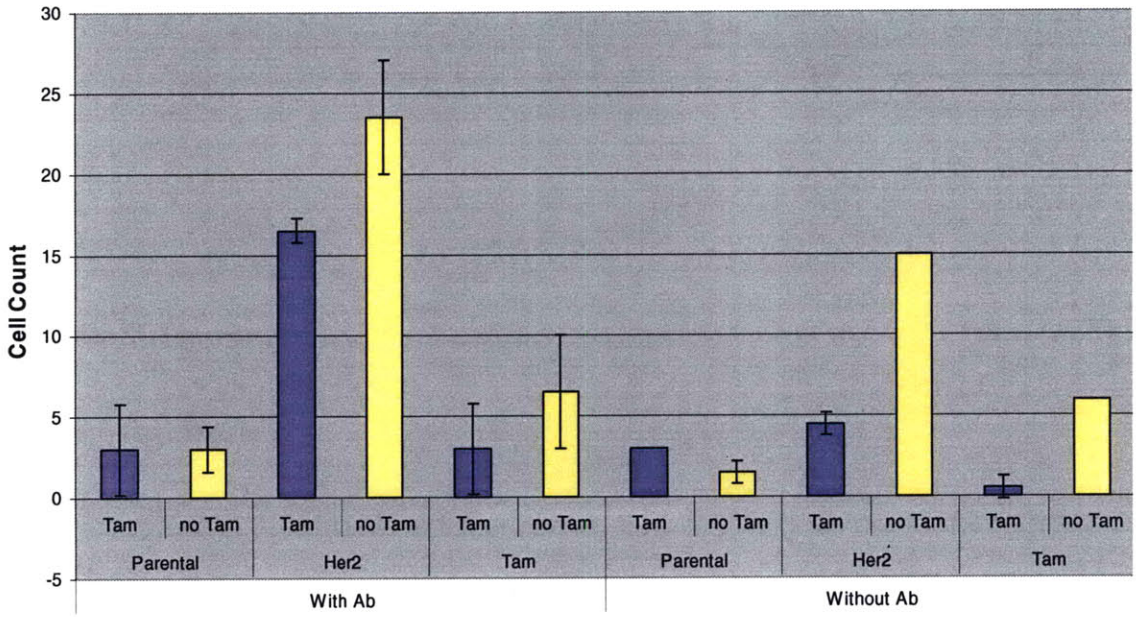


Figure 28. Invasion assay with E-cadherin antibody.

The effects of E-cadherin antibody treatment on the number of invasive cells were not as obvious as the effects seen in migration assays. Tamoxifen treatment did not increase the invasive ability of MCF7 cell lines. A major challenge in both migration and invasion assay with MCF7 cells is that the tightly-bound nature of MCF7 cells makes cells difficult to move despite treatment with E-cadherin antibody, which neutralizes E-cadherin function and loosens cell-cell contacts

mediated by E-cadherin.

II.4 Discussions

Tamoxifen resistance is an unmet medical need affecting a large number of patients with ER-positive breast cancer that have taken this therapeutic SERM. To address this problem we have applied an unbiased, mass spectrometry-based, quantitative proteomic analyses of protein phosphorylation to cell line models of Tamoxifen resistance. Our goal in this study was two-fold: first, to identify changes in the tyrosine phosphorylation-mediated signaling network caused by 4-OHT exposure; and second, to determine if this information would provide a potential therapeutic strategy to revert Tamoxifen resistance in different resistance models. Our analyses revealed that Tamoxifen resistant cell lines have the ability to maintain a high level of proliferation and activate anti-apoptotic signals in response to 4-OHT exposure, while the Tamoxifen sensitive cell line does not have the same capacity. In addition to the pro-survival (PI3K-Akt) and pro-proliferation (Erk) pathways, we also observed multiple hyperphosphorylated sites on adhesion and cytoskeletal proteins in MCF7-HER2, and to a lesser extent, MCF7-TAM cells in response to 4-OHT treatment. These molecules are likely controlled directly or indirectly by

increased activities of Src and FAK, but other pathway components cannot be ruled out. In fact, motif analysis of the amino acid sequences surrounding phosphorylation sites that increased phosphorylation in response to 4-OHT exposure in the MCF7-HER2 cells yielded a statistically significant enrichment of the Abl kinase motif. Abl is activated by Src and has been shown to promote invasion of aggressive breast cancer cells (Srinivasan and Plattner, 2006). Since dasatinib inhibits Src- and Abl-family kinases, with this compound it is not possible to distinguish whether Src or Abl is the principal driver modulating Tamoxifen resistance in the MCF7-HER2 and MCF7-TAM cell lines. However, given that this compound is currently used in the clinic as a second-line therapeutic for chronic myeloid leukemia, dasatinib may represent a clinically relevant option to revert Tamoxifen resistance regardless of the driving kinase.

It is also worth noting that since phosphorylation of catenin δ1, catenin δ2 and paxillin has been linked to detachment of cell-cell adhesions, these phosphorylation changes induced by 4-OHT may lead to increased migration and/or invasive potential (Hiscox et al., 2006). Moreover, since HER2 overexpression activates the cellular migration signaling network (Wolf-Yadlin et al., 2006), our

quantitative phosphorylation data indicates that Tamoxifen treatment of HER2 overexpressing cells may further increase the migratory and invasive potential of these cells, leading to increased aggressiveness.

The unbiased approach we have demonstrated here represents a valuable method to identify key signaling nodes at multiple layers in the complex network regulating therapeutic resistance. While perturbation of these nodes may not revert therapeutic resistance in every instance, here we have provided several examples in which inhibition of key nodes in the network has diminished Tamoxifen resistance to varying degrees, dependent on cell context. In the future, it might be possible to use selected phosphorylation sites (e.g. HER2 pY1248, catenin δ2 pY424 and pY499, or paxillin pY118) as markers to select the best therapeutic option (e.g. gefitinib, dasatinib, others) to revert or prevent Tamoxifen resistance, representing a potential application of personalized medicine for breast cancer patients.

REFERENCES

Beliakoff, J., Bagatell, R., Paine-Murrieta, G., Taylor, C.W., Lykkesfeldt, A.E., and Whitesell, L. (2003). Hormone-refractory breast cancer remains sensitive to the antitumor activity of heat shock protein 90 inhibitors. Clin Cancer Res *9*, 4961-4971.

Bellis, S.L., Miller, J.T., and Turner, C.E. (1995). Characterization of tyrosine phosphorylation of paxillin in vitro by focal adhesion kinase. J Biol Chem *270*, 17437-17441.

Benz, C.C., Scott, G.K., Sarup, J.C., Johnson, R.M., Tripathy, D., Coronado, E., Shepard, H.M., and Osborne, C.K. (1992). Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. Breast cancer research and treatment *24*, 85-95.

Bjornstrom, L., and Sjoberg, M. (2005). Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. Molecular endocrinology (Baltimore, Md *19*, 833-842.

Chaturvedi, L.S., Marsh, H.M., Shang, X., Zheng, Y., and Basson, M.D. (2007). Repetitive deformation activates focal adhesion kinase and ERK mitogenic signals in human Caco-2 intestinal epithelial cells through Src and Rac1. J Biol Chem *282*, 14-28.

Fan, Q.W., Knight, Z.A., Goldenberg, D.D., Yu, W., Mostov, K.E., Stokoe, D., Shokat,

K.M., and Weiss, W.A. (2006). A dual PI3 kinase/mTOR inhibitor reveals emergent efficacy in glioma. Cancer Cell *9*, 341-349.

Hanks, S.K., and Polte, T.R. (1997). Signaling through focal adhesion kinase. Bioessays *19*, 137-145.

Hiscox, S., Morgan, L., Green, T.P., Barrow, D., Gee, J., and Nicholson, R.I. (2006). Elevated Src activity promotes cellular invasion and motility in tamoxifen resistant breast cancer cells. Breast Cancer Res Treat *97*, 263-274.

Holmqvist, K., Cross, M., Riley, D., and Welsh, M. (2003). The Shb adaptor protein causes Src-dependent cell spreading and activation of focal adhesion kinase in murine brain endothelial cells. Cell Signal *15*, 171-179.

Huang, C., Ni, Y., Wang, T., Gao, Y., Haudenschild, C.C., and Zhan, X. (1997). Down-regulation of the filamentous actin cross-linking activity of cortactin by Src-mediated tyrosine phosphorylation. J Biol Chem *272*, 13911-13915.

Huang, P.H., Mukasa, A., Bonavia, R., Flynn, R.A., Brewer, Z.E., Cavenee, W.K., Furnari, F.B., and White, F.M. (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. Proceedings of the National Academy of Sciences of the United States of America *104*, 12867-12872.

Joughin, B.A., Naegle, K.M., Huang, P.H., Yaffe, M.B., Lauffenburger, D.A., and White, F.M. (2009). An integrated comparative phosphoproteomic and bioinformatic approach reveals a novel class of MPM-2 motifs upregulated in EGFRvIII-expressing glioblastoma cells. Molecular bioSystems *5*, 59-67.

Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., *et al.* (2009). Human Protein Reference Database--2009 update. Nucleic acids research *37*, D767-772.

Knowlden, J.M., Hutcheson, I.R., Jones, H.E., Madden, T., Gee, J.M., Harper, M.E., Barrow, D., Wakeling, A.E., and Nicholson, R.I. (2003). Elevated levels of epidermal growth factor receptor/c-erbB2 heterodimers mediate an autocrine growth regulatory pathway in tamoxifen-resistant MCF-7 cells. Endocrinology *144*, 1032-1044.

Lombardo, L.J., Lee, F.Y., Chen, P., Norris, D., Barrish, J.C., Behnia, K., Castaneda, S., Cornelius, L.A., Das, J., Doweyko, A.M., *et al.* (2004). Discovery of N-(2-chloro-6-methyl- phenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4- ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. J Med Chem *47*, 6658-6661.

Mariner, D.J., Anastasiadis, P., Keilhack, H., Bohmer, F.D., Wang, J., and Reynolds,

A.B. (2001). Identification of Src phosphorylation sites in the catenin p120ctn. J Biol Chem *276*, 28006-28013.

Moser, K., and White, F.M. (2006). Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. Journal of proteome research *5*, 98-104.

Neri, L.M., Borgatti, P., Capitani, S., and Martelli, A.M. (2002). The nuclear phosphoinositide 3-kinase/AKT pathway: a new second messenger system. Biochim Biophys Acta *1584*, 73-80.

O'Gorman, S., Fox, D.T., and Wahl, G.M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science *251*, 1351-1355.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic acids research *31*, 3635-3641.

Pedram, A., Razandi, M., Aitkenhead, M., Hughes, C.C., and Levin, E.R. (2002). Integration of the non-genomic and genomic actions of estrogen. Membrane-initiated signaling by steroid to transcription and cell biology. J Biol Chem *277*, 50768-50775.

Rexer, B.N., Engelman, J.A., and Arteaga, C.L. (2009). Overcoming resistance to tyrosine kinase inhibitors: lessons learned from cancer cells treated with EGFR

antagonists. Cell cycle (Georgetown, Tex *8*, 18-22.

Shou, J., Massarweh, S., Osborne, C.K., Wakeling, A.E., Ali, S., Weiss, H., and Schiff, R. (2004). Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer. J Natl Cancer Inst *96*, 926-935.

Srinivasan, D., and Plattner, R. (2006). Activation of Abl tyrosine kinases promotes invasion of aggressive breast cancer cells. Cancer research *66*, 5648-5655.

Whitesell, L., and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. Nat Rev Cancer *5*, 761-772.

Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H.D., Grantcharova, V., Lauffenburger, D.A., and White, F.M. (2006). Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. Mol Syst Biol *2*, 54.

Zhang, Y., Wolf-Yadlin, A., Ross, P.L., Pappin, D.J., Rush, J., Lauffenburger, D.A., and White, F.M. (2005). Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics *4*, 1240-1250.

Chapter III

Signal Transduction Analyses of Oncogenic Fusion Tyrosine Kinase EML4-ALK

Experiments performed by Hideshiro Saito-Benz[1] and Manabu Soda[2]

Experimental design by Hideshiro Saito-Benz[1], Manabu Soda[2], Hiroyuki Mano[2], and Forest M. White[1]

Author Affiliations

1. Department of Biological Engineering, David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge MA USA

2. Jichi Medical University, Tochigi Prefecture, Japan

III.1 SUMMARY

A subclass of non-small cell lung cancers (NSCLCs) expresses EML4-ALK chimeric oncogene. The EML4-ALK protein is a constitutively active tyrosine kinase with a transforming activity in NIH-3T3 fibroblast cells. Cell culture studies and transgenic mouse studies showed that EML4-ALK is a potential therapeutic target among a subclass of NSCLC patients. While the characteristics of EML4-ALK as a transforming gene and a constitutively active tyrosine kinase have been established, the signaling pathways activated by EML4-ALK remain unclear. In order to better understand EML4-ALK signaling networks, we used quantitative phosphoproteomic analyses by mass spectrometry to identify the tyrosine phosphorylation sites responsive to EML4-ALK expression. Results from the mass spec analyses revealed a number of highly phosphorylated proteins and their phosphorylation sites in 3T3 cells that are stably transfected with an EML4-ALK gene. Validations of EML4-ALK-specific phosphorylation sites by orthogonal methodologies showed that endogenous EML1 and EML4 are potential direct EML4-ALK substrates. Moreover, EML1 and EML4 may function as scaffolds in

order to recruit other EML4-ALK substrates to propagate the signals. Results obtained from this collaborative project shine light on novel EML4-ALK signaling pathways which are required to better understand its carcinogenic roles in NSCLCs.

III.2 INTRODUCTION

Lung cancer is one of the most difficult malignancies to cure among all cancers. Lung cancer is the leading cause of cancer related deaths. The number of new lung cancer cases is over 174,000 and the number of lung cancer related deaths is over 162,000 in the USA in 2005 (Jemal et al., 2006). Unfortunately, use of conventional chemotherapeutics rarely leads to cure mainly because of limited efficacy and because the majority of lung cancer patients are diagnosed at later stages of the disease (Schiller et al., 2002). Non-small cell lung cancers (NSCLCs) account for an approximately 80% of all lung cancer cases. Within NSCLC patients, a small subset of the patients carries EGFR activating mutations and such patients respond to the EGFR kinase inhibitor Iressa (Lynch et al., 2004), suggesting the mutated EGFR's critical role as a driver oncogene in this subset of patients (Paez et al., 2004). Improvements in the clinical outcomes of NSCLC patient populations without activating EGFR mutations are highly desired. Such improvements could be achieved by identifying promising therapeutic targets which are critical for tumor progression, as in the case of NSCLCs with mutated EGFR.

In order to identify novel driver oncogenes in NSCLCs, Soda et.al performed gene screenings with a retrovirus-mediated cDNA expression library created from a 62-year old NSCLC patient's specimen (Soda et al., 2007). As a result of infecting NIH-3T3 fibroblasts with the cDNA library and performing transformation assays, many transformed foci are formed, leading to the isolation and identification of cDNAs which caused foci formations. One of the isolated cDNAs was identified to contain an amino-terminal portion of human echinoderm microtubule-associated protein-like protein 4 (EML4) fused with a carboxyl-terminal portion of human anaplastic lymphoma kinase (ALK) including its tyrosine kinase domain. The biochemical assays analyzing this EML4-ALK fusion gene expressed in cell culture showed that the gene product is a transforming, constitutively active tyrosine kinase. EML4-ALK exists as a homodimer because the EML4 domain promotes dimerization. EML4-ALK interacts with microtubules in the cytoplasm, while wild-type ALK is a cell membrane embedded receptor tyrosine kinase. In addition to the original EML4-ALK, analysis of additional NSCLC specimens by RT-PCR with primers detecting EML4-ALK fusions identified additional EML4-ALK variants with distinctive gene fusion-points. The different variants of EML4-ALK all have the

ALK tyrosine kinase domain and gene fusion-points at similar locations as the original EML4-ALK gene, indicating that the variants are constitutively active tyrosine kinases. Although analyses of more specimens are required to generalize the observations, EML4-ALK$^+$ NSCLC specimens are absent for both EGFR mutations and KRAS mutations, suggesting the EML4-ALK$^+$ NSCLCs as a novel subclass within NSCLCs. Inhibition of EML4-ALK with ALK kinase inhibitor (WHI-P154) in the classic IL-3 dependent BaF3 transformation assays showed EML4-ALK's potential as a novel therapeutic target.

All of the characterizations described above were conducted in the NIH-3T3, HEK293, and BaF3 cell lines transfected with EML4-ALK. While such studies in cell lines are valuable, EML4-ALK's carcinogenic roles in the development and progression of NSCLCs remains unclear. In order to study EML4-ALK's carcinogenic role in a more physiological setting, tissue-specific transgenic mice that express EML4-ALK in lung epithelial cells were created. EML4-ALK knock-in mice developed a number of adenocarcinomas in both lungs of all transgenic mice, and mice died within six months after their birth (Soda et al., 2008). Importantly, oral administration of ALK kinase inhibitor 2,4-Pyrimidine-diamines derivative,

reduced tumor burden in all EML4-ALK transgenic mice, indicating that EML4-ALK induces carcinogenesis in the lungs of transgenic mice and an ALK inhibitor is effective in reducing the EML4-ALK induced carcinogenesis in the transgenic mouse model .

Although much work has been done to show the importance of EML4-ALK in lung carcinogenesis and its potential as a therapeutic target, the molecular signaling mechanisms by which EML4-ALK transforms normal cells and drives carcinogenesis remain unclear. In order to reveal potential carcinogenic mechanisms and activated signaling networks by EML4-ALK, we analyzed tyrosine phosphorylation events with a quantitative phosphoproteomic approach by mass spectrometry. Results from these analyses identified the tyrosine phosphorylation sites that are particularly responsive to EML4-ALK in 3T3 cells. Selected phosphorylation sites are validated by orthogonal methods in which EML4-ALK and each plasmid containing potential EML4-ALK target are co-expressed. Importantly, analyses of kinase-target interactions suggested the potential roles of endogenous EML1 and EML4 proteins as components of multi-phosphoprotein complexes. Results indicated that the multi-protein complex include a large number of

EML4-ALK substrates. The overall results reveal the important 'tip of iceberg' of potential pathways through which EML4-ALK initiates transformation in 3T3 fibroblasts, some of which are potentially conserved in NSCLCs.

III.3 MATERIALS and METHODS

III.3.1 Cell lines and reagents

The NIH-3T3 mouse fibroblast cells were stably transfected with retrovirus carrying 1) a vector with EML4-ALK L583M Kinase Dead Mutant (termed the KM which function as a negative control in this study), 2) a vector with wild type ALK, 3) a vector with EML4-ALK variant 1, or 4) a vector with EML4-ALK variant 3b in Mano lab in Jichi Medical University, Japan. Cells that were stably transfected were selected by the antibiotic selection. Upon arrival of frozen cell stocks to White Lab at MIT, cells are thawed and cultured in DMEM supplemented with 10% FBS until each plate reaches confluence. Cells are routinely subcultured at 1:10 when each plate reached confluence. Five frozen cell vial stocks of are prepared from one confluent 10cm cell culture plate and stored in a liquid nitrogen tank for future use. Fresh frozen stocks of cells are thawed at the beginning of each experiment in order to maintain consistencies across biological replicates. Cells are subcultured two rounds before experimental samples are collected. Cell lysate samples for

western blot analyses are collected from 10cm plates in RIPA buffer supplemented with PhoStop (Roche Applied Biosciences). Cell lysate samples for the phosphoproteomic analyses are collected from 15cm plates in 8M urea supplemented with 10mM sodium orthovanadate. Before collecting cell lysate samples, each plate were serum starved overnight in DMEM without FBS. Surprisingly, 3T3 cells expressing EML4-ALK variant 1 and variant 3b detached from plates a few minutes after serum starvation. Cells expressing the KM or wild type ALK did not detach from plates after serum starvation. In order to avoid cell detachments from plates, serum starvation step was not conducted in the real sample preparations. Instead of serum starvation, cells are gently rinsed with 37 degrees Celsius PBS in order to remove residual serum and subsequently lysed with appropriate lysis buffers. All the cell culture reagents are purchased from Invitrogen.

III.3.2 Peptide sample preparation.

For mass spectrometry analyses, cells were lysed in 8M urea buffer with

10mM sodium orthovanadate. Cell lysates were reduced with 10mM DTT and alkylated with 55mM iodoacetamide to block free thiols on cysteine side chains according to Zhang et. al (Zhang et al., 2005). The chemically modified cell lysate samples were subjected to tryptic digestion at a 1:100 ratio overnight and resulting tryptic peptides were desalted on a C18 reverse phase cartridge (Millipore). Peptides were eluted with 10mL of 25% acetonitrile/0.1% acetic acid, and 500µg peptide aliquots were lyopholized and stored at -80 degrees Celsius. Two biological replicate samples were prepared from two independent batches of cell cultures.

III.3.3 iTRAQ labeling of peptides and Peptide Immunoprecipitation

Peptide samples from 3T3 cell lines with four different vectors were chemically labeled with iTRAQ reagent according to Zhang et. al (Zhang et al., 2005). Four-plex iTRAQ reagents were used as follows: iTRAQ 114 for peptides from 3T3 cells with EML4-ALK Kinase Mutant (KM), iTRAQ 115 for peptides from 3T3 cells with wild-type ALK, iTRAQ 116 for peptides from 3T3 cells with EML4-ALK

variant 1, and iTRAQ117 for peptides from 3T3 cells with EML4-ALK variant 3b (Figure 1). After iTRAQ labeling, peptides were combined and subjected to anti-phosphotyrosine peptide immunoprecipitation as described (Zhang et al., 2005) with the following modifications: iTRAQ labeled peptides were incubated with 12µg of antiphosphotyrosine antibody (P-Tyr-100, (Cell Signaling Technology)) and 12µg of antiphosphorylation antibody 4G10 (Millipore) in 200ug of immunoprecipitation buffer for 8 hours at 4C, followed by incubation with 20ug of protein G Plus-agarose beads (Calbiochem) overnight.

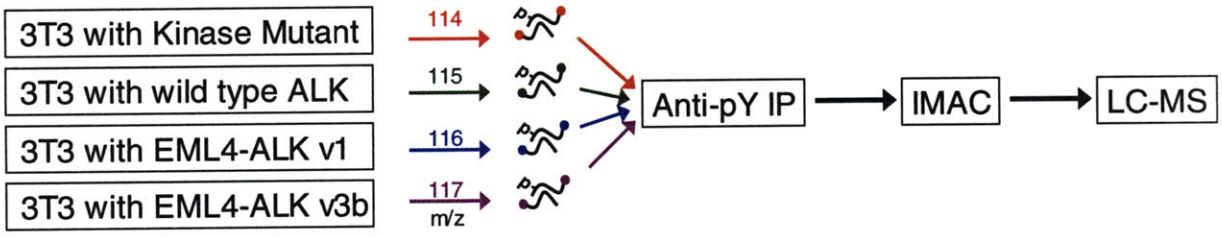


Figure 1. iTRAQ labeling of tryptic peptides from 3T3 cells with Kinase Mutant EML4-ALK L583M variant 1, wild type ALK, EML4-ALK variant 1, and EML4-ALK variant 3b

III.3.4 Immobilized metal affinity chromatography (IMAC) and Mass Spectrometry

Following anti-phosphotyrosine peptide immunoprecipitation, phosphopeptides were further enriched by immobilized metal affinity chromatography (IMAC) which removes non-specifically bound peptides from the immunoprecipitation (Moser and White, 2006). Peptides retained on an IMAC column were eluted with 250mM sodium phosphate at pH8.0 and analyzed by electrospray ionization liquid chromatography tandem MS on QSTAR XL Pro Mass Spectrometer (Applied Biosystems) (Zhang et al., 2005). MS/MS spectra were extracted, searched with Mascot (Matrix Science), and quantified with ProQuant (Applied Biosystems). Phosphorylation sites and peptide sequence assignments were validated by manual confirmation by assigning spectra of raw MS/MS data. Area under the peaks of iTRAQ reporter ions from phosphopeptides were normalized with values from the area under the peaks of iTRAQ reporter ion from non-phosphorylated peptides which is obtained from running combined iTRAQ peptides from supernatant of peptide immunoprecipitation on mass spectrometer. Total peptide quantification for each condition was measured by averaging peptide quantities from abundant house keeping genes that are expressed relatively constant levels, including but not limited to actin, tubulin, and HSP90.

III.3.5 Cell Titer Gro and WST-1 Assays

The KM, wild-type ALK, EML4-ALK v1, EML4-ALK v3b expressing 3T3 cell lines are thawed and passaged twice. Appropriate numbers of cells (1000, 3000, 10000) are seed onto 96 black well plates for cell titer gro assay, and clear 96 well plates for WST-1 assay. Seeded cells are undisrupted overnight and cell titer gro assay (Promega) and WST-1 (Roch Applied Bioscience) assays are performed according to manufactures protocols.

III.4 RESULTS

Two analytical replicates from two biological replicates (a total of four MS runs) are analyzed by quantitative mass spectrometry. Runs #1 and #2 are from the first biological replicate, and runs #3 and #4 are from the second biological replicate. Runs #1 and #4 had higher intensities of phosphopeptides and more phosphorylation sites detected than runs #2 and #3, respectively. For these reasons, run #1 and #4 were processed for quantification and manual spectra validation in order to obtain standard deviations and confirm identification of phosphorylation sites. From the compilation of run #1 and #4, 84 total tyrosine phosphorylation sites were identified and quantified. Out of 84 tyrosine phosphorylation sites, 55 are detected in both analyses (65%), while 29 sites are detected once in either run #1 or #4. A list of the 84 tyrosine phosphorylation sites, relative quantification across four conditions, standard deviations for 55 sites, and fold changes relative to the kinase mutant negative control are listed in table 2.

III.4.1 Wild type ALK and EML4-ALK activities measured by ALK pY1096

The phosphorylation at tyrosine 1096 of ALK (ALK pY1096) is in close proximity to its tyrosine kinase domain. ALK pY1096 is correlated to ALK kinase activities according to Phosphosite (www.phosphosite.org). iTRAQ reporter ion peaks from an ALK pY1096 peptide are shown in Figure 2. In the kinase dead mutant (the KM), ALK pY1096 was nearly absent (iTRAQ 114) indicating that the ALK activity is nearly absent in the KM sample. Relative to the KM, ALK pY1096 was 19.9-, 18.6-, and 32.5-fold higher in wild type ALK (iTRAQ 115), EML4-ALK v1 (iTRAQ 116), and EML4-ALK v3b (iTRAQ 117) respectively, indicating that ALK activities are higher in these three samples relative to the KM sample. In the wild type ALK sample, it is possible that FBS in the growth media contained ALK ligands for ligand-mediated activations. 3T3 cells transfected with wild type ALK grow preferably in close cell-cell contacts compared to the KM or EML4-ALK transfected 3T3 cells (data not shown). This specific phenotype indicates that 3T3 cells with wild type ALK have growth advantages when cells are in contact to each other. This is potentially because overexpressed wild type ALKs are activated by cell-cell contacts, either by cell surface ligands displayed on neighboring cells and/or other

proteins which is on outer surface of cell membrane. In the EML4-ALK v1 and v3b expressing 3T3 cell samples, ALK pY1096 are 18.6- and 32.5-fold higher than the KM. As 3T3 cells do not express wild type ALK, the ALK pY1096 is most likely from EML4-ALK transgenes. Before collecting cell lysate samples, each plate were serum starved overnight in DMEM without FBS. Surprisingly, 3T3 cells expressing EML4-ALK variant 1 and variant 3b detached from plates a few minutes after serum starvation. Cells expressing the KM or wild type ALK did not detach from plates after serum starvation. In order to avoid cell detachments from plates, serum starvation step was not conducted in the real sample preparations. Instead of serum starvation, cells are gently rinsed with 37 degrees Celsius PBS in order to remove residual serum and subsequently lysed with appropriate lysis buffers. Because serum starvation was not possible during sample preparation, it is probable that the FBS may activate signaling networks which may affect ALK pY1096, in addition to the previously proposed EML4-mediated oligomerization leading to EML4-ALK activation. Based on the strong increases on ALK pY1096 in wild type ALK, EML4-ALK v1, and v3 transfected 3T3 cells, we analyzed 83 other tyrosine phosphorylation sites in order to propose potential mechanisms of

EML4-ALK oncogenic signaling events.

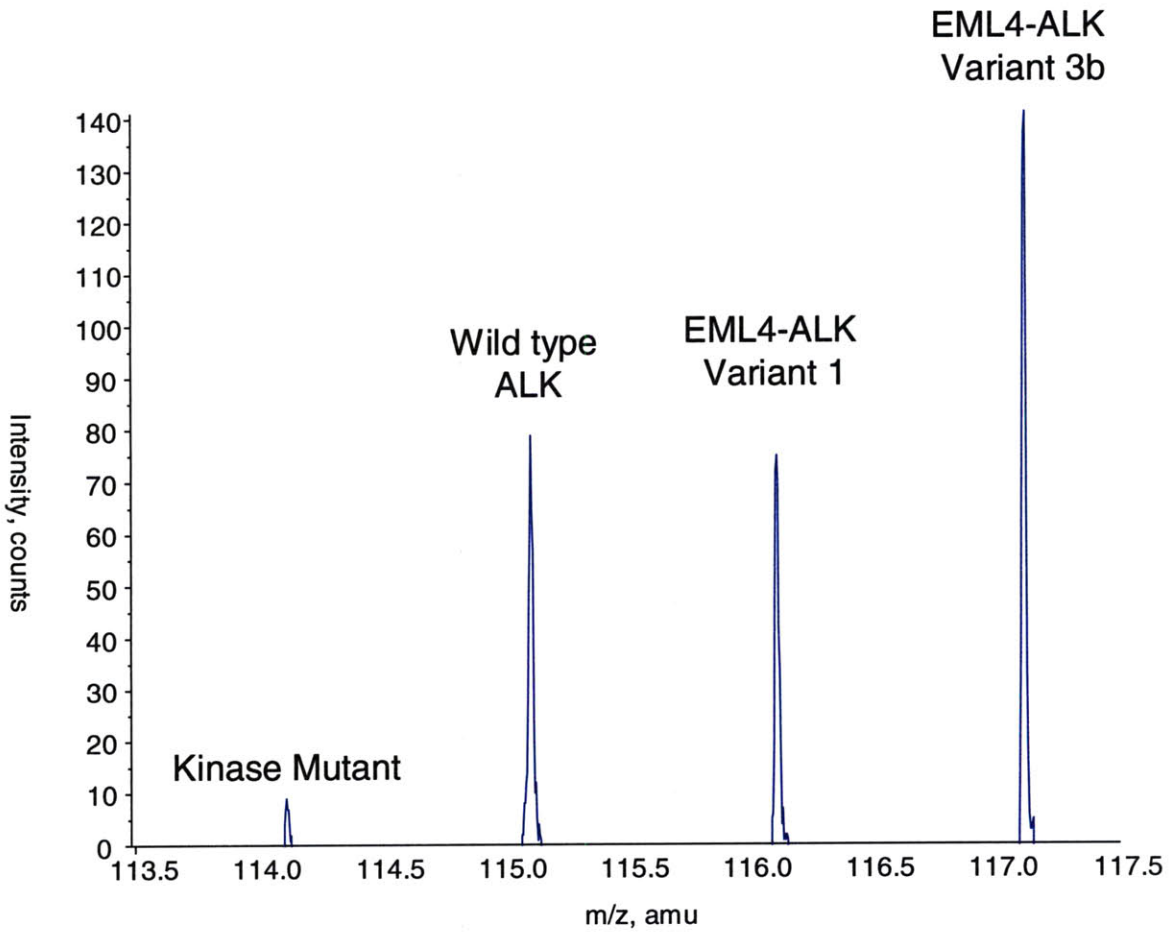


Figure 2. iTRAQ reporter ion peaks from an ALK pY1096 peptide. Representative iTRAQ reporter ion peaks from two biological replicate are shown.

III.4.2 Endogenous EML4 and EML1 are phosphorylated in EML4-ALK dependent manners

Several phosphorylation sites are found to be specifically responsive to EML4-ALK activities. Endogenous EML4 and EML1 are highly phosphorylated in EML4-ALK v1 and EML4-ALK v3b expressing 3T3 cells, but not in the KM or wild type ALK expressing 3T3 cells. Since 3T3 cells are mouse fibroblasts, we distinguished the slight amino acid differences in EML4 peptide sequences between human EML4-ALK transgenes and the endogenous EML4 protein, showing that the EML4 pY 226 detected is on the endogenous EML4 and not at the N-terminal part of EML4-ALK transgene. Endogenous EML4 pY226 is 11.9- and 18.0-fold higher in EML4-ALK v1 and EML4-ALK v3b expressing 3T3 cells than the KM. Endogenous EML1 is 35.2- and 51.2-fold higher in EML4-ALK v1 and EML4-ALK v3b expressing cells compared to the KM. In contrast, endogenous EML4 pY226 and EML1 pY186 are only 1.5- and 1.1-fold higher in ALK expressing 3T3 cells compared to the KM, showing that these two phosphorylation sites are increased by EML4-ALK but not by wild type AKL. Additionally, we detected endogenous EML1 pY396 which did not significantly change across four conditions, indicating that this phosphorylation site is not regulated by either ALK or EML4-ALK in this experimental system. IRS-2 has previously been proposed as a potential

downstream component of EML4-ALK in EML4-ALK^{+} positive lung tumor samples (Rikova et al., 2007). We found that IRS-2 pY776 is 2.1- and 7.5-fold higher in EML4-ALK v1 and EML4-ALK v3b samples respectively, whereas it remains unchanged in wild type ALK transfected 3T3 cells.

III.4.3. ALK and EML4-ALK dependent phosphorylation sites

In order to sort the 84 sites based on quantification patterns across four conditions, hierachical clustering was conducted by Spotfire (Figure 3).

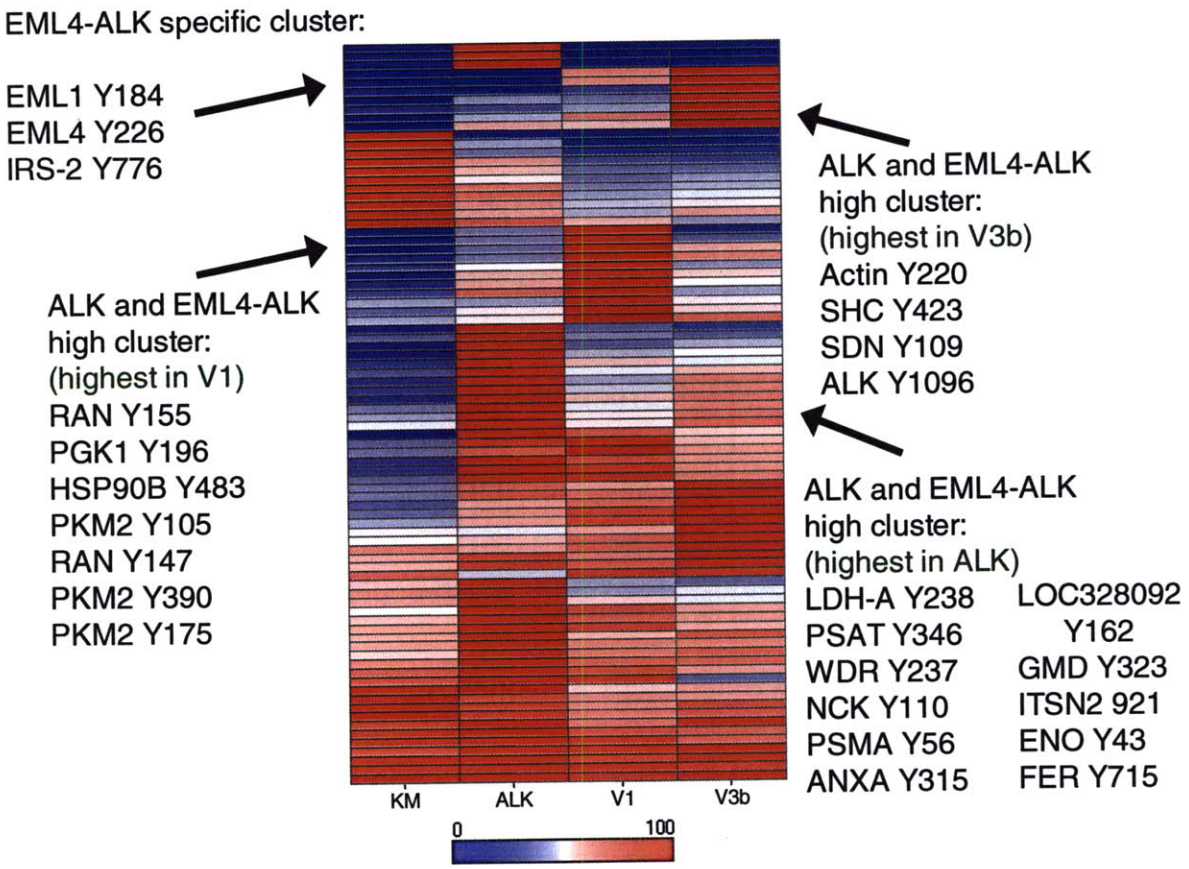


Figure 3. Hierachical clustering of 84 tyrosine phosphorylation sites. Relative quantification is normalized to the most abundant iTRAQ peak for each phosphorylation site. Scale is in percent. A cluster that responds to EML4-ALK, and three clusters that respond to both wild type ALK and EML4-ALK are shown.

A number of phosphorylation sites that increase in both wild type ALK and EML4-ALK dependent manners are clustered. These phosphorylation sites can be divided into three categories: A) the most responsive to EML4-ALK v3b, B) the most responsive to EML4-ALK v1, and C) the most responsive to wild type ALK.

As summarized in figure 3, cluster A includes actin pY220, SHC pY423, SDN pY109, and ALK pY1096. Cluster B includes RAN pY155, PGK pY196, HSP90B pY483, PKM2 pY105, RAN pY147, PKM2 pY390, OKM2 pY175. Cluster C includes LDH-A pY238, PSAT pY346, WDR pY237, NCK pY110, PSMA pY56, ANXA pY315, LOC328092 pY162, GMD pY323, ITSN2 pY921, ENO pY43, and FER pY715. In cluster A, ALK pY1096 is an indicative of ALK activity. 3T3 cells expressing EML4-ALK v3b have the highest ALK pY1096 which could be due to differences in transgene expression levels and/or due to higher ALK activity of EML4-ALK v3b relative to wild type ALK and EML4 ALK v1. SHC pY423 regulates association of adaptor molecules to induce their interactions with Grb2 (Faisal et al., 2004; Sasaoka et al., 2001; Zhang et al., 2002). SND1, which is alternatively known as a p105 co-activator, is a co-activator of STAT6, bridging STAT6 with RNA Polymerase II in order to initiate transcription at STAT6 regulated genes. In category B, as EML4-ALK localize in microtubules, highly phosphorylated RAN pY155 and pY147 indicate regulations of microtubule networks by RAN, independently of its roles in the transport of macromolecules between nucleus and cytosol. HSP90B pY483 is highly phosphorylated in both EML4-ALK and ALK dependent manners (4.9-, 14.0-,

and 8.4-fold in wild type ALK, EML4-ALK v1 and EML4-ALK v3b expressing 3T3 cells compared to the KM). HSP90B is one of the HSP90 isoforms and is an abundant chaperon protein. PKM2 pY105, pY390, pY175 are highly phosphorylated in both ALK and EML4-ALK dependent manners. PKM2 pY390 is particularly highly phosphorylated (13.2-, 22.3-, and 12.4-fold higher in wild type ALK, EML4-ALK v1 and EML4-ALK v3b expressing 3T3 cells compared to the KM). PKM2 has recently been shown to be a phosphotyrosine binding protein, even though it lacks known phosphotyrosine binding SH2 or PTB domains (Christofk et al., 2008). In category C, a number of vital enzymes in metabolic pathways, including LDH-A, PSAT, and ENO are highly phosphorylated. LDH-A catalyzes the reaction between pyruvate and lactate in an NADH dependent manner. ENO converts 2-phosphoglycerate to phosphoenolpyruvate during glycolysis. PSAT converts 2-phosphohydroxypruvate into serine during amino acid biosynthesis. Overall, many enzymes are highly phosphorylated in both wild type ALK and EML4-ALK dependent manners. As metabolic enzymes are known to be tightly regulated by post-translational modification, it is possible that signals from ALK and EML4-ALK modulate activities of enzymes leading to abnormal metabolic pathway

levels.

III.4.4 ALK dependend phosphorylation sites

Because ALK is a receptor tyrosine kinase signaling from the cell membrane, we predicted that localization of the EML4-ALK in microtubules results in at least partial loss of normal ALK signaling pathways. As seen in Figure 4, ALK high cluster has three phosphorylation sites that are exclusively high in wild type ALK transfected 3T3 samples.

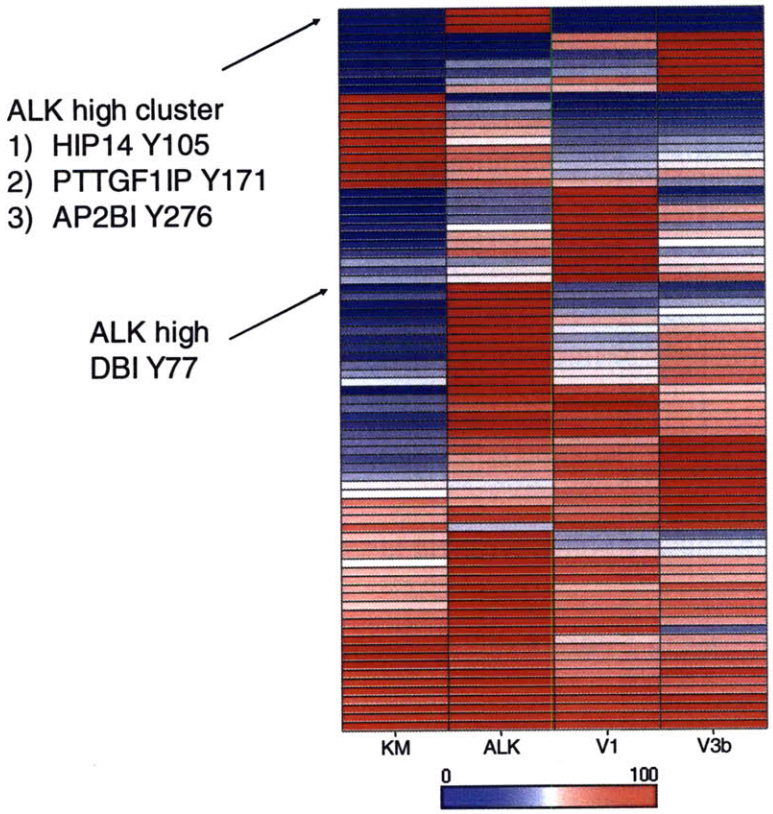


Figure 4. Cluster of ALK specific phosphorylation site and one phosphorylation site that is highly ALK responsive and weakly EML4-ALK responsive is shown.

These sites are HIP pY105, PTTGF1IP pY171, and AP2BI pY276. Intriguingly, common characteristics of these three ALK dependent phosphorylation sites are that these proteins are membrane associated (Stowers and Isacoff, 2007). PTTFG1IP is a type 1a integral membrane protein and contains the tetrapeptide YXRF, a motif observed in proteins internalized by coated pit-mediated endocytosis.

AP2B1 is a part of the AP2 coat assembly protein complex that link clathrin to many cell surface receptors. Another ALK dependent phosphorylation site is DBI pY77. DBI is a protein known to be involved in membrane phospholipid turnovers. Lack of these four phosphorylation sites in EML4-ALK transfected 3T3 cells indicates that EML4-ALK localization in proximity to microtubules results in a loss of interactions between EML4-ALK to the proteins localized cell membrane. It also indicates that EML4-ALK is not only a constitutively active tyrosine kinase, but also is a distinctive kinase from wild type ALK in having different downstream pathway components.

III.4.5 Phosphorylation sites that are lost in ALK and EML4-ALK expressing 3T3 cells

There are three proteins which are hypophosphorylated in response to wild type ALK or EML4-ALK expression in 3T3 cells. These three sites are Eph A3/A4/A5 pY779, Tensin pY1480, and Caveolin-1 pY14 (Figure 5).

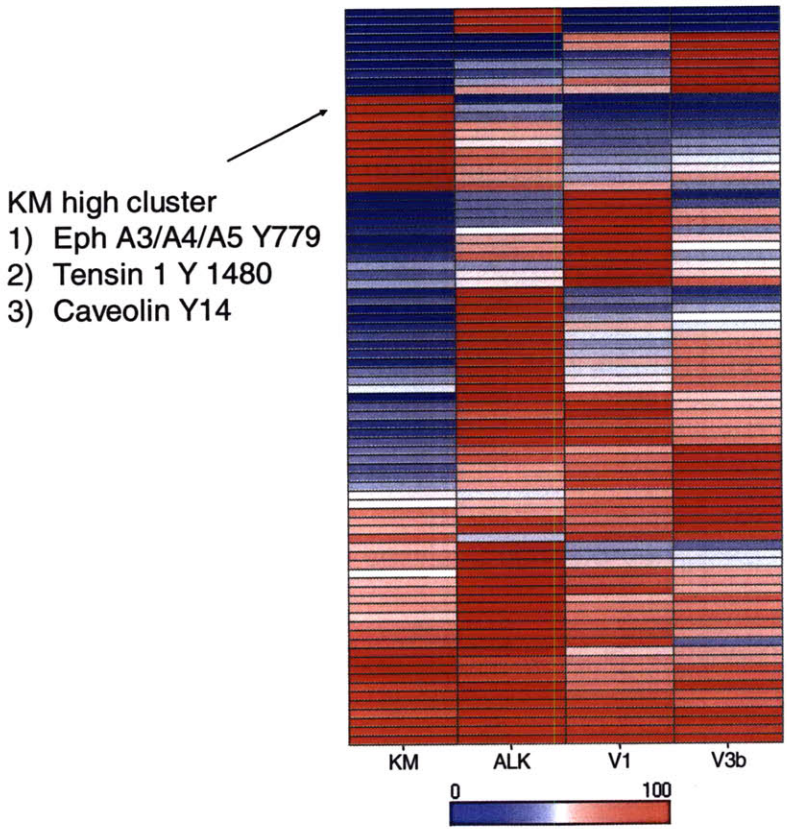


Figure 5. A cluster of phosphorylation sites that are hypophosphorylated in wild type ALK and EML4-ALK expressing 3T3 cells.

Caveolin-1 mediates integrin-regulated membrane domain internalization, which is a mechanism used to shut down growth factor receptor signaling. Decrease in Caveolin-1 pY14 leads to constitutive activation of growth regulatory pathways, and often results in anchorage independence (del Pozo et al., 2005). Decreased Caveolin-1 phosphorylation in both wild type ALK and EML4-ALK expressing 3T3

cells indicates that growth regulatory pathways are deregulated, leading to fast growth and anchorage independence. There is no previous literature for Tensin 1 pY1480, and biological functions of the phosphorylation site remain unknown.

III.4.6 Interactions between endogenous EML4 and EML1 with EML4-ALK

After completion of phosphoproteomic analyses, validations of endogenous EML4 and EML1 phosphorylation sites as EML4-ALK substrates were performed in the Mano lab at Jichi Medical University. To confirm interactions between the kinase and its substrates, MYC-tagged EML4-ALK was transfected to 3T3 cells along with FLAG-tagged human EML1 or human EML4 and anti-MYC immunoprecipitations followed by anti-FLAG blots are performed (Figure 6).

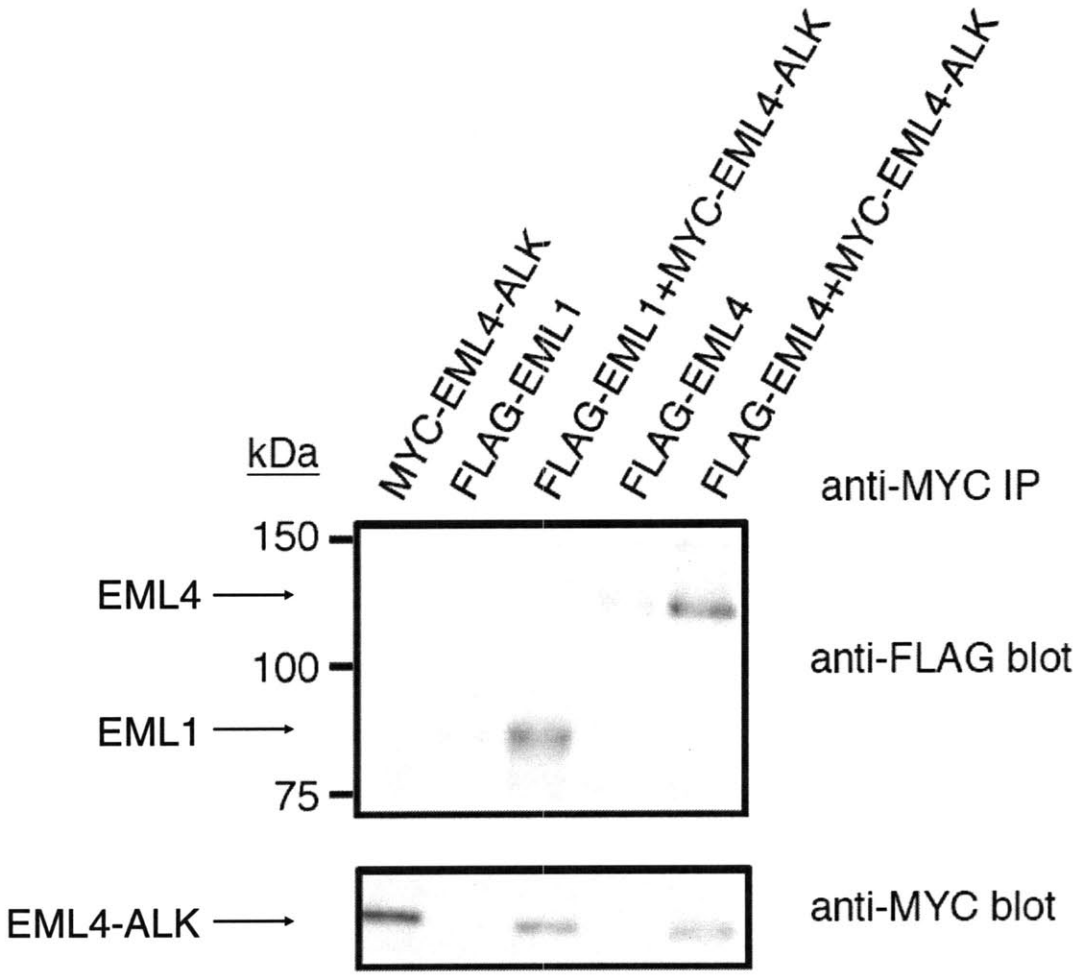


Figure 6. Interactions between EML1 or EML4 and EML4-ALK.

Based on this result, human EML1 and human EML4 interact with EML4-ALK. At this point, it was uncertain if EML1 and EML4 phosphorylation sites are required for the interaction between EML1 or EML4 and EML4-ALK. In order to test the function of the EML1 and EML4 phosphorylation site, the tyrosine phosphorylation

sites on both proteins were mutated to phenylalanine (YF), and the same experiments were repeated (Figure 7). YF mutant EML1 and EML4 could maintain interaction with EML4-ALK, although to lesser extents. The weaker interactions between YF mutant EML1 or YF mutant EML4 and EML4-ALK indicate that a loss of tyrosine phosphorylation site might be the cause of weaker interactions, although further experiments with loading controls are necessary to make a conclusion

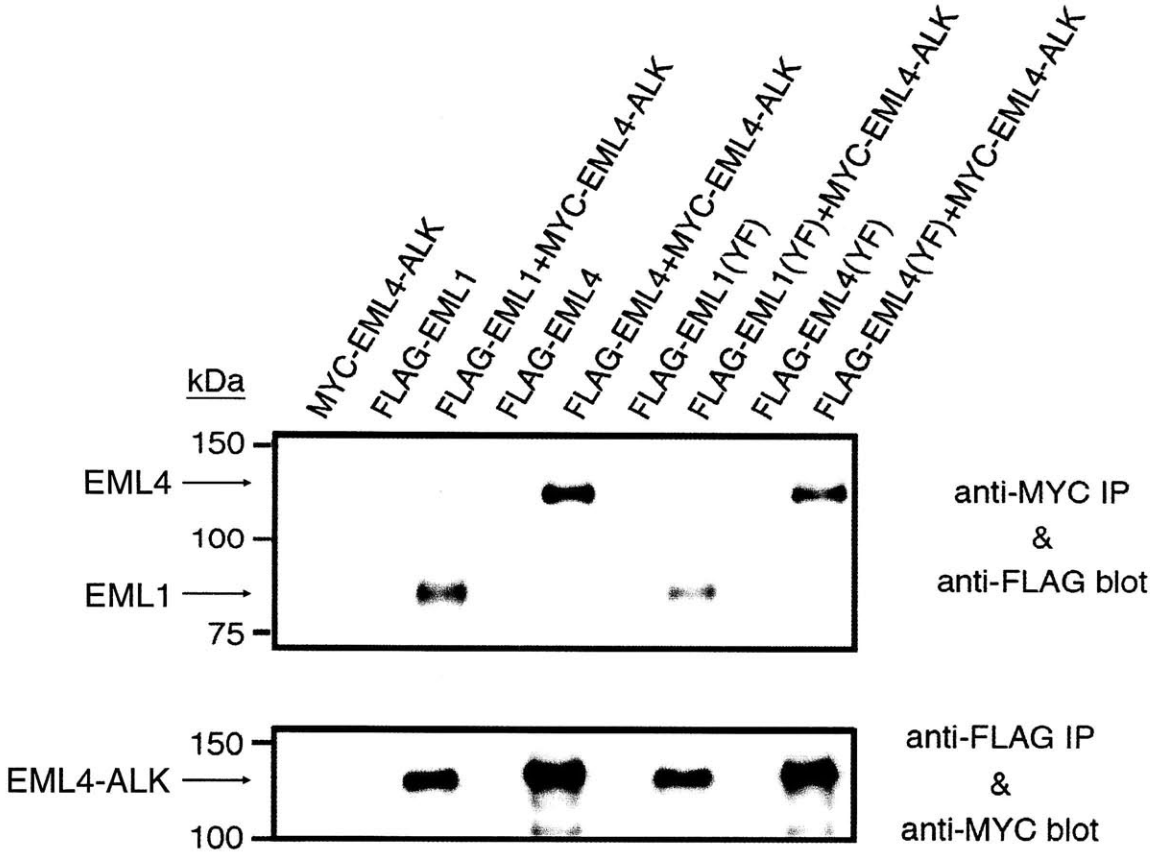


Figure 7. Interactions between EML1 or EML4 and EML4-ALK.

Although EML4 proteins are known to homodimerize, it remained unclear if EML4 can heterodimerize with other EML isoforms such as EML1. In order to test if EML1 and EML4 interact each other and if the interaction is EML4-ALK dependent, FLAG-tagged EML4 and HA-tagged EML1 were expressed in 3T3 cells with or without EML4-ALK, and immunoprecipitations followed by immuno blots were performed (Figure 8).

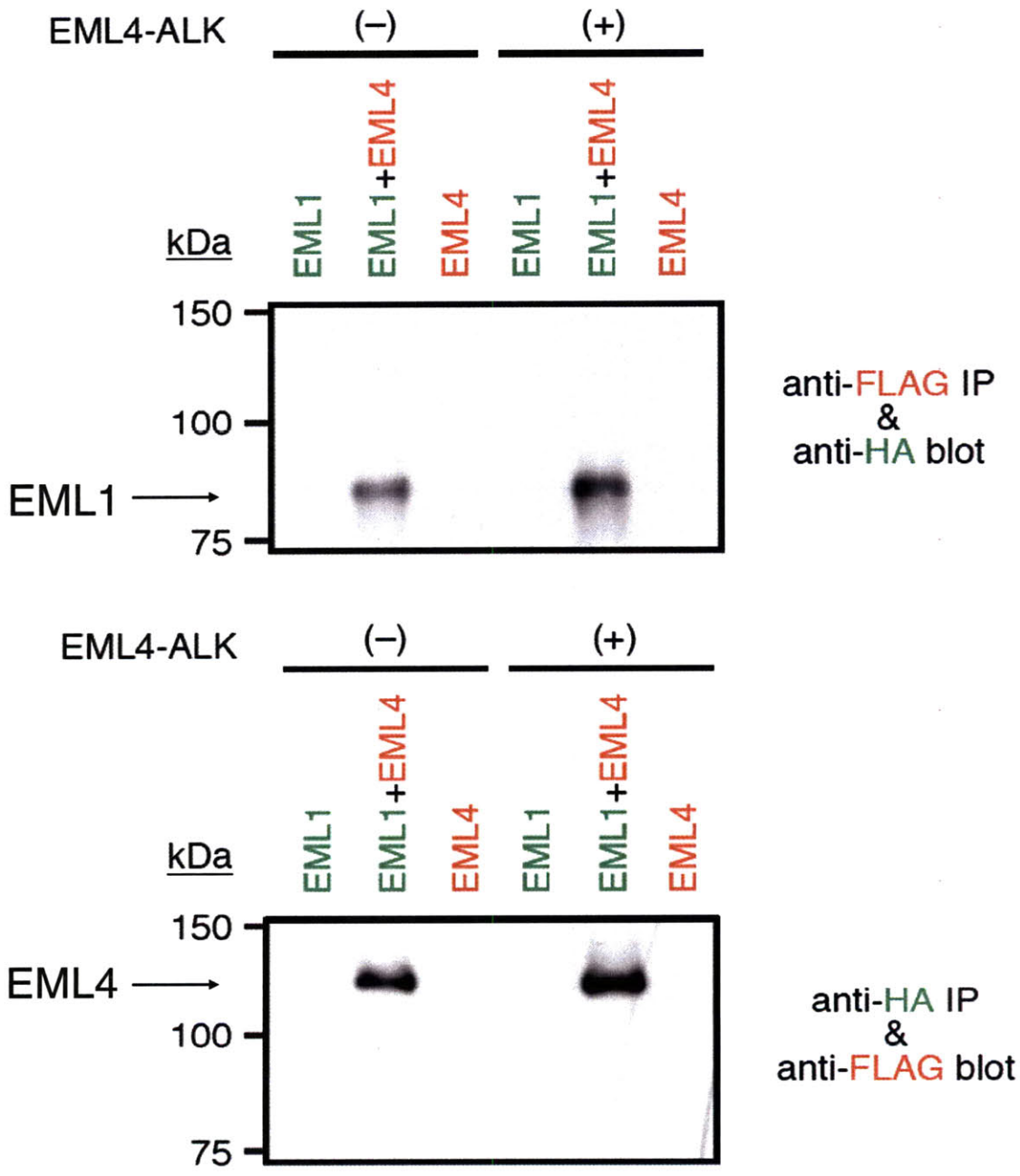


Figure 8. Interactions between EML1 and EML4 occurs in the absence of EML4-ALK.

The anti-FLAG immunoprecipitation and anti-HA blot, and anti-HA

immunoprecipitation and anti-FLAG blot confirm interaction of EML1 and EML4. The interaction was present in the absence of EML4-ALK, while the interaction seemed to increase in the presence of EML4-ALK, although the loading controls for this experiment need to be shown. It remains to be seen if the interactions between EML1 and EML4 will be weaker when EML4-ALK dependent phosphorylation sites are mutated to phenylalanine.

III.4.7 EML1 and EML4 interact with tyrosine phosphorylated proteins in EML4-ALK dependent manner.

Although EML1 and EML4 interact with EML4-ALK, the functional consequences of such interactions remained unclear. In addition, it is possible that there are additional EML4-ALK substrates that are part of EML4-ALK/EML1 and EML4-ALK/EML4 complexes. In order to analyze EML4-ALK/EML1 and EML4-ALK/EML4 complexes, FLAG-tagged EML4, EML1, YF mutated EML4 or YF mutated EML1 were expressed in 3T3 cells in the presence or absence of MYC-tagged EML4-ALK. Samples are subjected to anti-FLAG

immunoprecipitation followed by anti-pY blot (Figure 9).

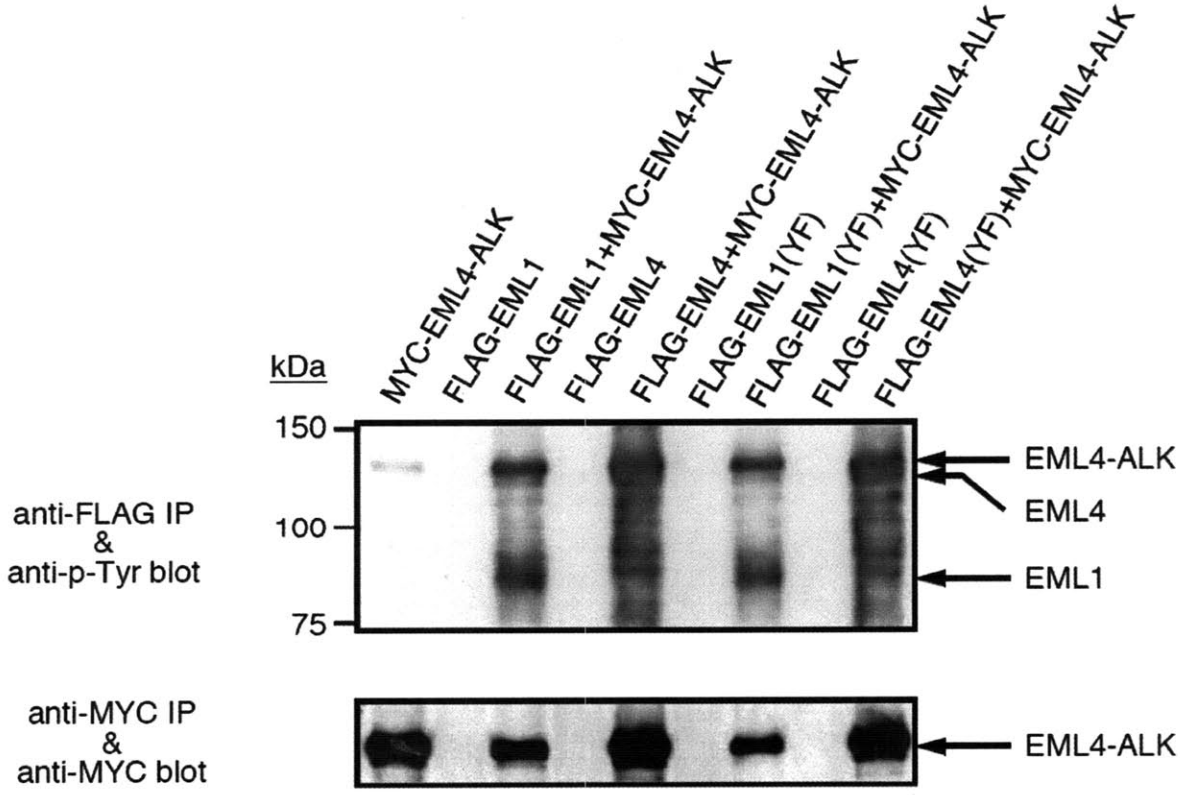


Figure 9. EML1 and EML4 interact with tyrosine phosphorylated proteins in the presence of EML4-ALK.

Wild type EML1 and EML4 interact with many tyrosine phosphorylated proteins in the presence of, but not in the absence of, EML4-ALK. YF mutant EML and YF mutated EML4 can interact with tyrosine phosphorylated proteins in the presence of EML4-ALK. These results suggest that 1) tyrosine phosphorylation of EML1 and EML4 associated proteins occur EML4-ALK dependent manners, and 2) mutant

EML1 and EML4 maintains partial functions which could be due to the presence of unidentified EML4-ALK substrate tyrosines on EML1 and EML4.

III.4.8 Analysis of phosphorylation sites on EML1 and EML4

In order to identify additional phosphorylation sites on EML1 and EML4 that are responsive to EML4-ALK expression, 3T3 cells are co-transfected with EML1 or EML4 with EML4-ALK. EML1 or EML4 are affinity purified, and run on SDS-PAGE in order to isolate the EML1 or EML4 containing bands for mass spec analyses (Figure 10).

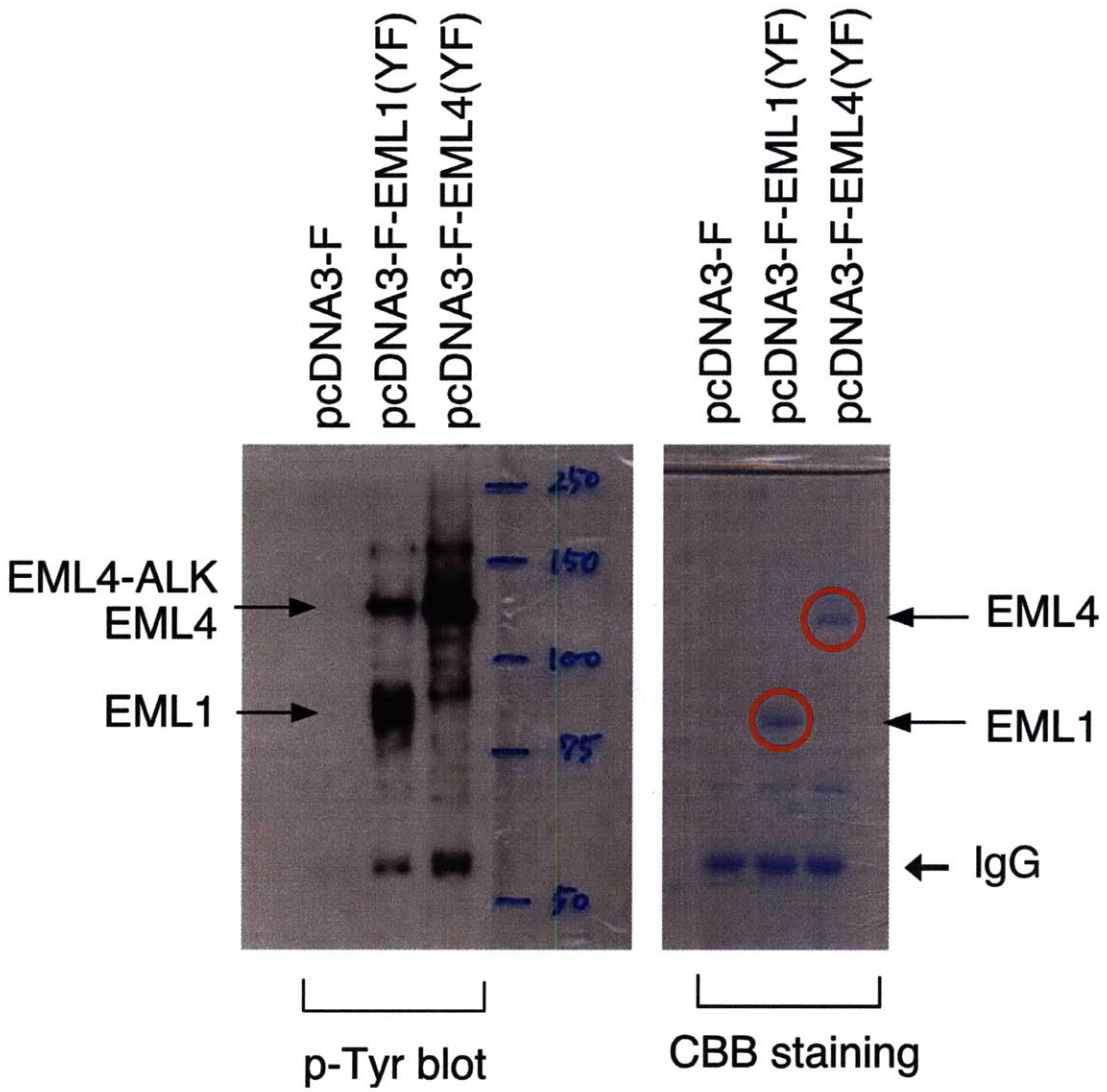


Figure 10. EML1 and EML4 are immunoprecipitated and blotted with 4G10 phosphotyrosine antibody (left) or stained with the commersie blue (right). The commercie blue stained bands containing EML1 and EML4 are isolated and shipped on dry ice to MIT.

Gel bands containing EML1 or EML4 were isolated and in-gel digested in order to

perform IMAC-LC-MS/MS analyses. Three peptide aliquots are prepared from a single gel band, and two analytical replicate analyses were performed on EML1 and EML4 gel bands (total four analyses). We quality-controlled each step of IMAC-LC-MS/MS, and analyzed both IMAC-retained and IMAC-flowthrough peptides in each analysis. Although a number of EML1 and EML4 non-phosphorylated peptides are identified in IMAC flow-through, phosphorylated peptides were not identified in IMAC-retained peptides.

III.4.9 Analysis of energy production and metabolic activity in EML4-ALK expressing 3T3 cells.

We have hypothesized that EML4-ALK expressing cells maintain a high level of energy productions and metabolisms. To test this hypothesis, the KM, EML4-ALK v1, and EML4-ALK v3b expressing 3T3 cells were seeded onto 96 well plates to conduct a cell titer grow assay and a WST-1 assay (Figure 11).

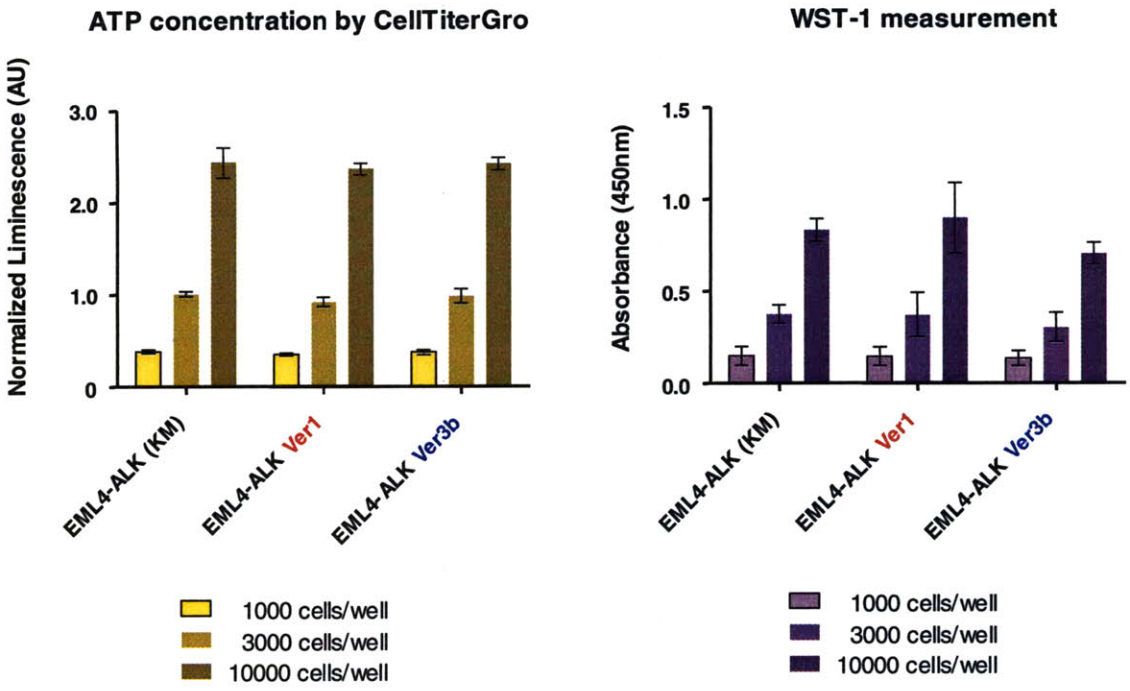


Figure 11. Measurement of metabolic activities of 3T3 cells with EML4-ALK Kinase Mutant, EML4-ALK variant 1, and EML4-ALK variant 3b.

The cell titer gro assay measures total levels of ATP in cells, and WST-1 measures indirectly a level of NADPH dependent enzymatic activities. Contrary to our expectations, neither EML4-ALK v1 nor v3b increased the readings from Cell titer gro or WST-1 assay compared to the KM control. It is possible that both assays were not the appropriate measurements of altered metabolic pathways, and/or experimental conditions were not suitable to observe potential metabolic differences between the KM and EML4-ALK expressing 3T3 cells. It is not certain if the

incrased phosphorylation levels are related to high ATP production and/or metabolism. The highly phosphorylated enzymes potentially regulate amino acid production, nucleic acid synthesis, and lipid biosynthesis which are regarded as highly active in cancerous cells.

III.5 Discussions

EML4-ALK is a fusion tyrosine kinase that is constitutively active and transforming in 3T3 cells. In order to understand its potential transforming and carcinogenic mechanisms, we employed the quantitative phosphoproteomic analyses by mass spectrometry to reveal EML4-ALK induced tyrosine phosphorylation signaling networks. From the phosphoproteomic analyses, 84 phosphorylation sites are identified. A number of phosphorylation sites are found to be responsive to EML4-ALK expression.

An important positive control is the ALK pY1096 levels, which are indicative of ALK tyrosine kinase activities. EML4-ALK L586M kinase dead mutant (the KM) expressing 3T3 cells had a minimal level of ALK pY1096, which shows that this mutation inactivates ALK. In wild type ALK transfected 3T3 cells, ALK pY1096 is as high as in the EML4-ALK v1 transfected 3T3 cells, indicating that ALK kinase activities in each condition might be similar. This suggests that the signaling networks initiated from different cellular locations might be significantly different. Wild type ALK and fusion EML4-ALK localize at different locations and have

distinctive activation mechanisms. Wild type ALK is a membrane bound receptor tyrosine kinase which is activated by unknown mechanisms. The ligands for ALK are still unknown, and activation mechanisms are predicted to be ligand-mediated homodimerization followed by autophosphorylation. EML4-ALK is a constitutively active homodimeric cytosolic tyrosine kinase localized to microtubules. Protein populations around each kinase's environments are different, and the differences in signaling networks were revealed between similar ALK kinase activities based on different localizations. One question that remains to be solved is the potential activation mechanisms of wild type ALK. The ligands for ALK have not been discovered yet, and the mechanisms by which ALK tyrosine kinase domain phosphorylates and fully activate its kinase activities remains unclear. Based on the growth of wild type ALK transfected 3T3 cells that grow preferentially in contact with neighboring cells, it is possible that ALK activities are induced by cell-cell contacts. It remains unclear what proteins expressed on the surface of plasma membrane have affinities toward extracellular domains of ALK. EML4-ALK v1 and EMl4-ALK v3b are constitutively active kinases and v3b is predicted to have stronger tyrosine kinase activities (Mano, H, Personal communication).

EML4-ALK v3b has the highest ALK pY1096 among all conditions which indicates that 3T3 cells expressing EMl4-ALK v3b has the highest ALK kinase activities, whether it is due to the empirical expression level biases or innate differences between EML4-ALK v1 and v3b.

One important question from all signaling analyses of kinase transfected cell lines is if the kinase expression level is physiologically relevant because expression levels can significantly alter substrate specificities (Jones et al., 2006). In our study design, we focused on detection of tyrosine phosphorylation sites from 3T3 mouse fibroblast cells that were transformed by EML4-ALK expression. This is the classic model of transformation assay that was used to demonstrate the transforming activities of BCR-ABL (Daley et al., 1987). 3T3 cell focus formation assay with EML4-ALK is further complimented by IL3 dependent BaF3 transformation assays, which is the gold standard for *in vivo* screening process for BCR-ABL kinase inhibitors (Daley and Baltimore, 1988). Similar to BCR-ABL, EML4-ALK is shown to transform IL3 dependent BaF3 cell lines and growth could be inhibited by increasing doses of ALK kinase inhibitors. Based on these characteristics, we hypothesized that analyses of EML4-ALK tyrosine kinase

signaling network in 3T3 cells are relevant for understanding its molecular mechanisms.

One of the important findings from this study is that endogenous EML1 and EML4 are phosphorylated in EML4-ALK dependent manners, while both the KM and wild type ALK does not affect EML1 and EML4 phosphorylation levels. These results were confirmed in 3T3 cells and 293 cells by co-transfection of EML4-ALK and human EML1 or EML4 followed by anti-EML1 and EML4 immunoprecipitation and anti-phosphotyrosine blots with 4G10 (Mano, H, Personal communications). Therefore, EML1 and EML4 phosphorylation by EML4-ALK are not 3T3 cells specific events, and the phosphorylation occurs at least in two different cell lines.

EML1 and EML4 physically interact with EML4-ALK, which is indicative of a kinase-substrate relationship. The interaction continued when the EML4-ALK dependent tyrosine phosphorylation sites on EML1 and EML4 were mutated to phenylalanine, indicating that these tyrosine phosphorylation sites are not required for the interaction. Analyses of YF mutants of EML1 and EML4 in the presence of EML4-ALK with EML1 and EML4 immunoprecipitation followed by anti-phosphotyrosine blots suggested the presence of additional EML4-ALK

dependent tyrosine phosphorylation sites (Mano H, personal communications). It is possible that there are unidentified EML4-ALK dependent tyrosine phosphorylation sites on EML1 and EML4 that compliments a loss of a single phosphorylation site resulting from the YF mutation. Furthermore, EML1 and EML4 interact with each other in 3T3 cells. While experiments with proper loading controls are required to make a conclusion, presence of EML4-ALK seems to increase the interaction between EML1 and EML4, while YF mutation of EML1 and EML4 decreases such interaction.

Wild type ALK and EML4-ALK transfected cells have increased phosphorylation levels on many important cellular enzymes in metabolic pathways (Figure 12).

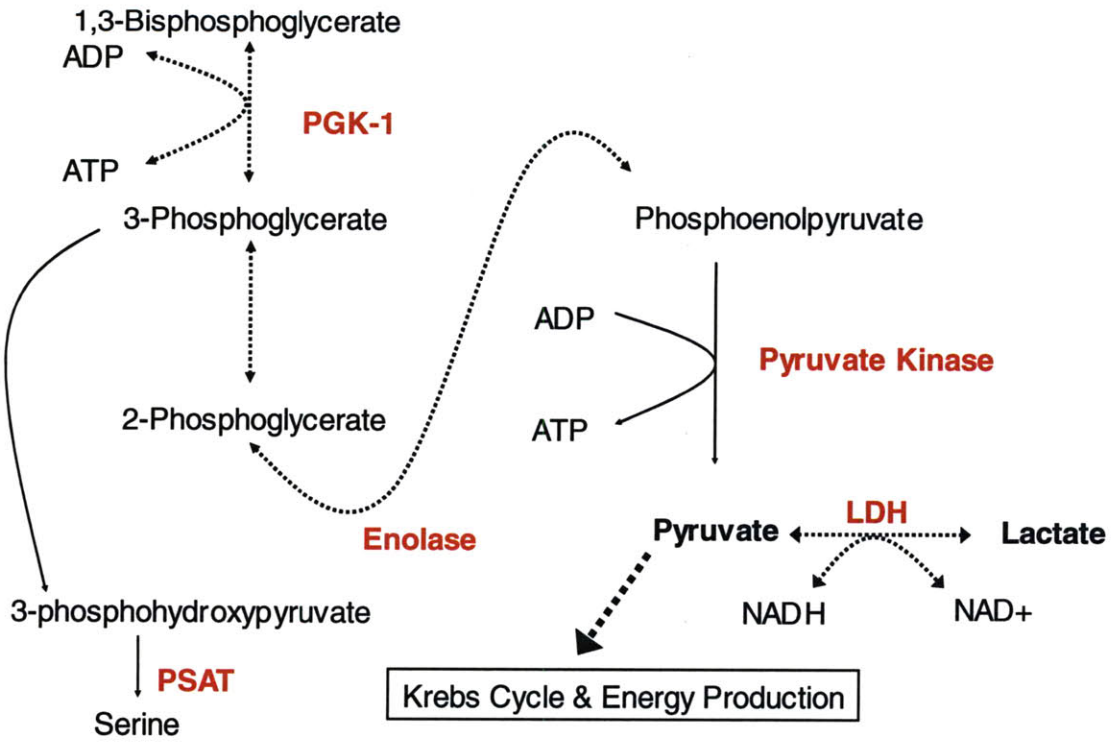


Figure 12. Enzymes in metabolic pathways are highly phosphorylated in ALK and EML4-ALK expressing 3T3 cells. Hyperphosphorylated enzymes are highlighted in red.

Notably, PKM2 pY390 is highly phosphoryatled in an EML4-ALK dependent manner. While functional insights of many phosphorylation sites on metabolic enzymes are still missing, the involvement of metabolic enzymes in cancer homeostasis is starting to become clear. ENO and LDH-A were highly phosphorylated in an EML4-ALK dependent manner. While functional outcomes of each phosphorylation event remain unclear, it indicates that a kinase with high activity

can potentially reprogram metabolic pathways in order to maintain the high energy production and biosynthetic demands. Highly proliferative cells such as cancerous cells requires production of basic cellular materials, such as amino acids, nucleic acids, and lipids. The highly phosphorylated enzymes potentially regulate amino acid production, nucleic acid synthesis, and lipid biosynthesis in order for cancerous cells to continuously divide at high rates.

An important finding from the validation process of phosphoproteomic data is that EML1 and EML4 seem to interact with many tyrosine phosphorylated proteins in the presence of EML4-ALK as visualized in Figure 13.

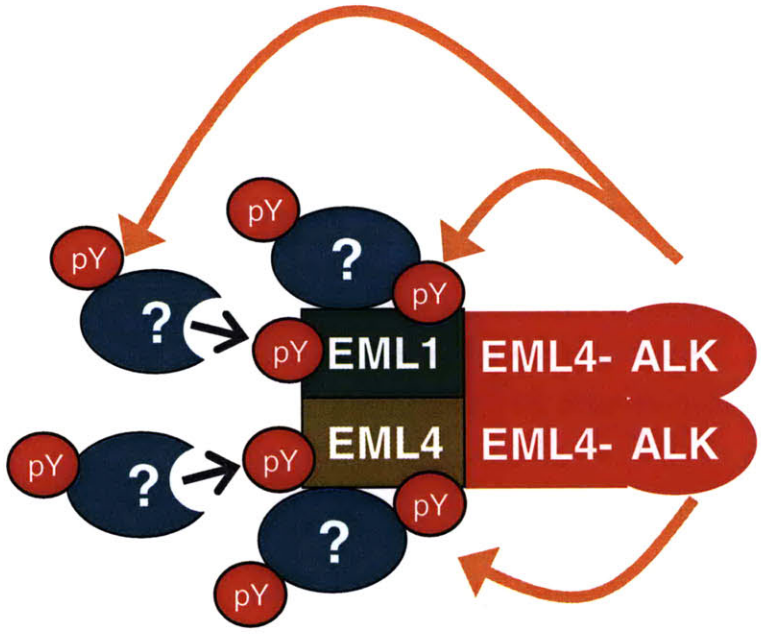


Figure 13. Schematic diagram of EML4 and EML1 mediated protein phosphorylation by EML4-ALK. EML1 and EML4 are phosphorylated by EML4-ALK. Phosphorylated EML1 and EML4 function as adaptors or scaffolds, recruiting proteins. Proteins bound on EML1 and EML4 are phosphorylated by EML4-ALK

While it remains unclear if EML4-ALK interacts with EML1 directly or if endogenous EML4 acts as a bridges between EML1 and EML4-ALK, it seem that EML4-ALK at least partially exists as a part of multi-protein complexes including EML1 and EML4 at microtubules. It is probable that many SH2- and PTB-containing proteins are recruited by EML1 and EML4 tyrosine phosphorylation sites in EML4-ALK dependent manners. These recruited proteins may be phosphoryated by EML4-ALK due to its proximity to the kinase. It is speculative at this point, and further experiments are required to confirm that EML4-ALK phosphorylates many proteins that are part of a large protein complex.

In conclusion, the studies conducted in this collaborative project shines lights on the mechanisms of EML4-ALK signal transductions. EML4-ALK is a promising therapeutic target for NSCLC patients who are diagnosed as an EML4-ALK$^+$ subclass. A treatment against the fusion tyrosine kinase BCR-ABL has been shown to be successful from the history of the BCR-ABL discovery and

Gleevec development in Chronic Mylogenous Leukemia patients. Fusion genes are previously thought to play pathogenic roles in hematologic malignanancies, but not in solid tumors. EML4-ALK is the first example of a transforming fusion tyrosine kinase in a solid tumor. In order to develop therapeutic compounds effectively, a better understanding of EML4-ALK signaling mechanisms is vital at early stages of therapeutic developments. Through a series of experiments in this collaborative project, we have revealed the tip of an iceberg of EML4-ALK signaling mechanisms. The summary of proposed mechanisms of EML4-ALK signaling mechanisms is shown in Figure 14. We hope that further discoveries will follow for the benefit of NCSLC patients.

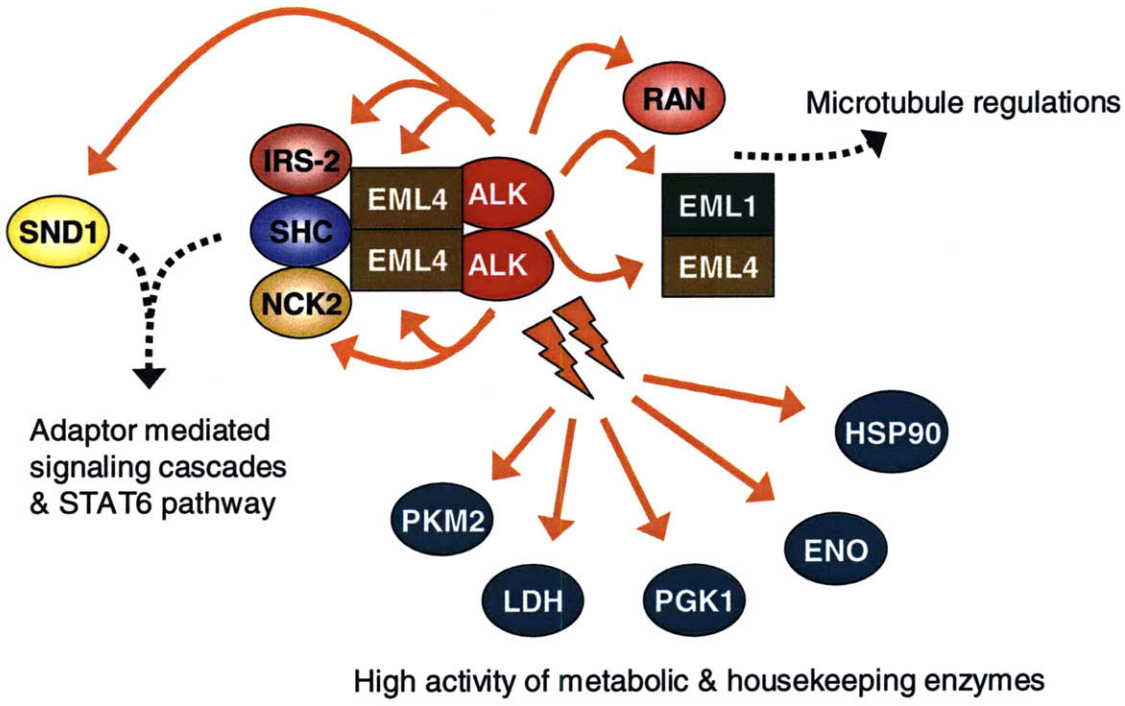


Figure 14. Tip of an iceberg of EML4-ALK oncogenic signaling mechanisms.

References

Christofk, H.R., Vander Heiden, M.G., Wu, N., Asara, J.M., and Cantley, L.C. (2008). Pyruvate kinase M2 is a phosphotyrosine-binding protein. Nature *452*, 181-186.

Daley, G.Q., and Baltimore, D. (1988). Transformation of an interleukin 3-dependent hematopoietic cell line by the chronic myelogenous leukemia-specific P210bcr/abl protein. Proceedings of the National Academy of Sciences of the United States of America *85*, 9312-9316.

Daley, G.Q., McLaughlin, J., Witte, O.N., and Baltimore, D. (1987). The CML-specific P210 bcr/abl protein, unlike v-abl, does not transform NIH/3T3 fibroblasts. Science (New York, NY *237*, 532-535.

del Pozo, M.A., Balasubramanian, N., Alderson, N.B., Kiosses, W.B., Grande-Garcia, A., Anderson, R.G., and Schwartz, M.A. (2005). Phospho-caveolin-1 mediates integrin-regulated membrane domain internalization. Nature cell biology *7*, 901-908.

Faisal, A., Kleiner, S., and Nagamine, Y. (2004). Non-redundant role of Shc in Erk activation by cytoskeletal reorganization. The Journal of biological chemistry *279*, 3202-3211.

Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., Smigal, C., and Thun, M.J. (2006). Cancer statistics, 2006. CA: a cancer journal for clinicians *56*, 106-130.

Jones, R.B., Gordus, A., Krall, J.A., and MacBeath, G. (2006). A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature *439*, 168-174.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., *et al.* (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. The New England journal of medicine *350*, 2129-2139.

Moser, K., and White, F.M. (2006). Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. Journal of proteome research *5*, 98-104.

Paez, J.G., Janne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., *et al.* (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science (New York, NY *304*, 1497-1500.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., *et al.* (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell *131*, 1190-1203.

Sasaoka, T., Ishiki, M., Wada, T., Hori, H., Hirai, H., Haruta, T., Ishihara, H., and Kobayashi, M. (2001). Tyrosine phosphorylation-dependent and -independent role of Shc in the regulation of IGF-1-induced mitogenesis and glycogen synthesis. Endocrinology *142*, 5226-5235.

Schiller, J.H., Harrington, D., Belani, C.P., Langer, C., Sandler, A., Krook, J., Zhu, J., and Johnson, D.H. (2002). Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. The New England journal of medicine *346*, 92-98.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., *et al.* (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561-566.

Soda, M., Takada, S., Takeuchi, K., Choi, Y.L., Enomoto, M., Ueno, T., Haruta, H., Hamada, T., Yamashita, Y., Ishikawa, Y., *et al.* (2008). A mouse model for EML4-ALK-positive lung cancer. Proceedings of the National Academy of Sciences of the United States of America *105*, 19893-19897.

Stowers, R.S., and Isacoff, E.Y. (2007). Drosophila huntingtin-interacting protein 14 is a presynaptic protein required for photoreceptor synaptic transmission and expression of the palmitoylated proteins synaptosome-associated protein 25 and cysteine string protein. J Neurosci *27*, 12874-12883.

Zhang, L., Camerini, V., Bender, T.P., and Ravichandran, K.S. (2002). A nonredundant role for the adapter protein Shc in thymic T cell development. Nature immunology *3*, 749-755.

Zhang, Y., Wolf-Yadlin, A., Ross, P.L., Pappin, D.J., Rush, J., Lauffenburger, D.A., and White, F.M. (2005). Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol Cell Proteomics *4*, 1240-1250.

IV. Conclusions

IV. CONCLUSION

Through a series of connected experiments in this thesis, I have identified a molecular mechanism by which breast cancer cells adapt to pressures from anti-estrogen compounds. Tamoxifen resistance mechanisms are complex, likely multifactoral phenomena. Discovering key differences between Tamoxifen sensitive and Tamoxifen resistant tumors is critical for better clinical practice. I have addressed this problem by applying novel phosphoproteomic technology to understand differences in tyrosine phosphorylation events. The thesis work described in the previous chapters can be divided into three major sections, which reflect the specific aims of the thesis.

In the first section, I have derived Tamoxifen resistant cell lines using previously described methods. Long term low dose Tamoxifen exposure to the Tamoxifen sensitive MCF7 cell line indeed lead to the development of an acquired Tamoxifen resistant cell line (MCF7-TAM). In addition, stable overexpression of HER2 tyrsosine kinase in MCF7 cells (MCF7-HER2) lead to increased growth rates as well as maintenance of growth in the presence of Tamoxifen.

MCF7-TAM is a population of Tamoxifen resistant cells that adapted to pressures from exposure to Tamoxifen. MCF7-TAM's growth profile is essentially similar to parental MCF7 in the absence of Tamoxifen, whereas complete resistance against Tamoxifen treatment is observed in the presence of Tamoxifen treatment. Interestingly, MCF7-TAM maintains a similar ER expression level as parental MCF7. Despite the similar ER expression level, growth of MCF7-TAM becomes less dependent on hydrophobic factors present in serum, including estrogen.

MCF7-HER2 cells are populations of MCF7 cells that are expressing approximately 40-fold more HER2 protein. According to previous work by CC. Benz et al, HER2 overexpressing MCF7 acquired Tamoxifen resistance *in vitro* and *in vivo*. The *in vitro* observation was reproduced independently in this study. MCF7-HER2 cells remain estrogen dependent for their growth based on the growth profile conducted in estrogen-depleted media.

Both MCF7-TAM and MCF7-HER2 in this study have distinctive growth profiles. MCF7-TAM is minimally affected by Tamoxifen treatment, whereas MCF7-HER2 is greatly affected by Tamoxifen treatment compared to MCF7-TAM, yet sustains a higher level of growth compared to parental MCF7. They represent

different models of Tamoxifen resistant populations. Therefore the common adaption/resistance mechanisms are potentially promising therapeutic targets for large numbers of breast cancer patients because interventions against such targets might be beneficial to two different populations of Tamoxifen resistant patients.

In the second section, I profiled the differences in the tyrosine phosphorylation events of Tamoxifen sensitive and two Tamoxifen resistant cell lines. One hundred twenty tyrosine phosphorylation sites were identified and quantified from multiple biological replicates of the phosphoproteomic analyses. Quantification patterns on each phosphorylation site were compared with growth profiles of each cell line in order to find relationships between tyrosine phosphorylation patterns and phenotypic outcomes. We started analyzing quantification patterns of tyrosine phosphorylation sites on proteins with known proliferative/anti-apoptotic functions and key downstream components of HER2 signaling networks. For example, PI3K regulatory subunit 2 pY605 (PI3K R2 pY605) decreased approximately 50% in parental MCF7 in response to Tamoxifen treatment. Interestingly, PI3K R2 pY605 increased approximately by 100% and 70% in MCF7-TAM and MCF7-HER2 respectively in response to Tamoxifen

treatment. Further analyses suggested correlations between quantification patterns of PI3K R2 pY605 and growth profiles of each cell lines. While it remains to be validated, PI3K R2 pY605 is likely to interact with a SH2 subunit of its binding partner PI3K regulatory subunit 1. Such an interaction recruits a PI3K p110 catalytic subunit to activate the trimeric type I PI3K. From the breast tumor analysis, we were unable to detect PI3K pY605, therefore it remains to be seen if the same activation occurs in primary tumors.

To measure activities of PI3K pathways in Tamoxifen sensitive and Tamoxifen resistant cell lines, we assessed the level of downstream Akt activity by measuring pS473 Akt with ELISA. Intriguingly, Tamoxifen resistant cells had higher pS473 Akt compared to Tamoxifen sensitive cells in response to Tamoxifen treatment. In addition to ELISA, western blot analyses showed that pS473 Akt is increased in Tamoxifen resistant cells in response to Tamoxifen treatment (K. Pant, personal communication). MCF7-TAM cells had significantly higher pS473 Akt than we predicted from PI3K R2 pY605 quantifications. This is potentially due to inactivation of PTEN due to S-nitrosation in MCF7-TAM cells, leading to deregulated PIP_2 and PIP_3 levels as discussed in the NOS inhibitor section.

Besides the PI3K-Akt pathway, we saw correlations between ERK1/2 phosphorylation patterns and growth profiles of each cell line. ERK1/2 phosphorylation levels were similar between parental MCF7 and MCF7-TAM in the absence of Tamoxifen, whereas ERK1/2 phosphorylaiton levels were 50% higher in the MCF7-HER2 than two other cell lines. Growth profiles were in good agreement with ERK1/2 phosphorylation patterns in that parental MCF7 and MCF7-TAM grew at similar rates, while MCF7-HER2 grew more rapidly than either cell line. In response to Tamoxifen treatment, parental MCF7 and MCF7-HER2 showed decreased growth rates which corresponds to 50% decreased ERK1/2 phosphorylation sites in each cell line. MCF7-TAM cells maintained comparable growth rates both in the absence or presence of Tamoxifen, and such growth rates match to unchanged phosphorylation levels on ERK1/2 in MCF7-TAM in response to Tamoxifen. In addition to ERK1/2 phosphorylation levels from cell line models, analyses from primary tumors showed that tyrosine phosphorylation events on four MAPKs including ERK1/2 were greatly altered in the recurrent tumor that I analyzed. The results obtained from primary tumors were similar to results obtained from cell line studies in that ERK1/2 phosphorylation sites were high in Tamoxifen treated

samples.

In addition to PI3K and MAPK pathways, we identified the Tamoxifen-induced activations of Src/FAK tyrosine kinases and their downstream pathways. Such activations were unique to Tamoxifen resistant cells and were not detected in Tamoxifen sensitive cells. Src and FAK phosphorylation levels were increased by 50% and 70% in the MCF7-HER2 cells. Numbers of known and predicted Src/FAK kinase substrates were found to be hyperphosphorylated in Tamoxifen resistant cells in response to Tamoxifen treatment. In addition to Src and FAK pathway activation, bioinformatic analysis of the hyperphosphorylated peptide sequences from Tamoxifen resistant cells predicted the enrichment of Abl target domains, indicating that Abl kinase may also be activated in Tamoxifen resistant cells. From the results of primary breast tumor analyses, Src family kinase and FAK were consistently hyperphosphorylated in the Tamoxifen-treated recurrent tumor sample. Because few tyrosine phosphorylation sites on known and predicted Src/FAK/Abl kinase substrates were detected in breast tumor analyses, it remains unclear if the activities of Src/FAK/Abl kinases lead to the hyperphosphorylation of their substrates. Analyses of additional human primary

tumors may strengthen such hypotheses.

In the third section, I conducted small-molecule inhibition experiments to revert Tamoxifen resistance by inhibiting identified therapeutic candidates from phosphoproteomic analyses. In order to identify therapeutic targets, we generated hypotheses from phosphoproteomic data. Candidate targets were hyperphosphorylated kinases that may play important roles in Tamoxifen resistant cell lines, including PI3K, ERK1/2, and Src/Abl.

In order to inhibit PI3K pathways, we used two direct PI3K inhibitors (LY294002 and PI103). Both MCF7-HER2 and MCF7-TAM responded to the PI3K inhibitors in combination with Tamoxifen. Compared to the positive control used in the inhibition study (Iressa, which was previously shown to revert Tamoxifen resistance), the extent of growth inhibition by PI3K inhibitors were similar to an extent of growth inhibition of the positive control on MCF7-HER2, while PI3K inhibition did not achieve the desired degree of inhibition in MCF7-TAM.

A second strategy was to inhibit MEK by PD98059 to decrease downstream ERK1/2 activities. MEK inhibition was effective in MCF7-TAM cell lines, indicating that MCF7-TAM needs ERK1/2 activity to maintain growth in the presence of

Tamoxifen. MEK inhibition was less effective in MCF7-HER2 cells than in MCF7-TAM cells, indicating that MCF7-TAM cells rely on ERK1/2 pathway for their survival, while MCF7-HER2 relies more on the PI3K pathway for growth and survival in the presence of Tamoxifen. Interestingly, Iressa failed to inhibit the growth of MCF7-TAM cells, indicating ERK1/2 activation in MCF7-TAM cells is independent of EGFR activity. As MCF7-TAM increase iNOS expression in response to Tamoxifen treatment and NOS inhibitor treatment inhibits NO production, S-nitrosated phosphatases, and decrease viable cell numbers, it is possible that deactivated phosphatases for ERK1/2 are the potential cause of the ERK 1/2 activation in MCF7-TAM cells.

I analyzed phosphoproteomic data further in order to identify signatures from tyrosine phosphorylations that may explain Tamoxifen resistance mechanisms in MCF7-HER2 and MCF7-TAM cells. Based on comparisons of phosphorylation sites, protein functions, and bioinformatic analyses of the phosphoproteomic data, we identified activation signatures of Src/Abl pathways exclusively in Tamoxifen resistant cells. To probe the functions of Src/Abl pathway activations in Tamoxifen resistance, we employed the dual-specificity Src/Abl kinase inhibitor Dasatinib

(Sprycel), which is an FDA-approved second line therapeutic compound for chronic mylogenous leukemia. The effects of combinatorial treatment with Dasatinib and Tamoxifen were compared with positive controls (HSP90 inhibitors, Geldanamycin and 17AAG). Combinatorial treatments with Dasatinib and Tamoxifen were effective in suppressing the growth of both MCF7-HER2 and MCF7-TAM to a similar degree to the positive controls. Inhibition of Abl specifically by Gleevec would be an interesting experiment because such an experiment would distinguish if the activated kinase is either Src, Abl, or both. We are planning to conduct this experiment in the near future.

These results suggest that different Tamoxifen resistant cell line models utilize multiple pathways that cells rely on for their growth. Identification of common pathways and key nodes were successful for targeted therapeutic treatments in Tamoxifen resistant cellular models. In the case of MCF7-HER2, PI3K and Src/Abl pathways may play key roles for growth sustenance. For MCF7-TAM, ERK1/2 and Src/Abl pathways were up-regulated, and MEK and Src/Abl were the optimal therapeutic candidates for this model. In order to find optimal targets for both models of Tamoxifen resistance, Src/Abl pathways were

tested and found as effective candidates, as use of Dasatinib was shown to be the most useful option among all compounds we tested in this thesis work. Based on these in vitro results, I believe that Dasatinib in combination with Tamoxifen would be potentially more effective than Tamoxifen alone in the clinic. Although further in vivo analyses in animal models are required, such a combinatorial treatment may improve the prognosis of hormone receptor positive breast cancer patients.

V. FUTURE PERSPECTIVES

V. FUTURE PERSPECTIVES

I have discussed the current state of endocrine therapy to combat breast cancer, and identified potential therapeutic targets to revert Tamoxifen resistance by application of phosphoproteomic analyses of tyrosine phosphorylation signaling networks. While I believe such findings will eventually benefit breast cancer patients, there are additional important problems which are beyond the scope of this thesis project. I would like to discuss some of these problems in this section.

Breast cancer and heterogeneity

Breast tumors are heterogeneous populations which consist of many cancerous cells with different expression levels of various driver oncogenes. In modern clinical practice, patients diagnosed with breast cancer are histologicaly categorized into well defined categories of breast cancers based on the expression levels of diagnostic marker genes such as ER/PR hormone receptors and HER2. Variation in HER2 and ER/PR expression levels within a single breast tumor cell

population is one example of heterogeneity. All cancerous cells within the breast tumor tissue do not express both diagnostic marker genes even if a patient is classified as $HER2^+/ER^+$ by immunohistochemistry. Within such tumors, $HER2^-/ER^+$ cells are potentially an optimal target of Tamoxifen treatment, while $HER2^+/ER^+$ tumor cells are growth-inhibited by Tamoxifen, while $ER^-/HER2^+$ tumor cells may not be affected. If the majority of the tumor cell population consists of the optimal targets of a therapeutic regimen, patients are expected to benefit the most. However, when a patient's immunohistochemical diagnostic results are on a borderline between $HER2^+/ER^+$, $HER2^-/ER^+$, $ER^+/HER2^-$ status, Tamoxifen treatment may not be effective due to the heterogeneous responses that each cell type likely to have. Such conditions are, to a certain extent, analogous to experimental conditions where parental MCF7 and MCF7-HER2 are mixed together and exposed to Tamoxifen. While it remains difficult to analyze any co-cultured cell lines, such an experiment may simulate the effects of Tamoxifen treatment to heterogeneous cell populations with different expression levels of molecular markers. Due to complex molecular interactions within the cell-cell microenvironment, it is incredibly challenging to conduct such experiments in cell

culture systems and analyze the effect for such heterogeneous conditions in two-dimentional, two cell-population cell culture conditions. However, it is important to note that such co-cultured system may be useful for understanding complex adaptation mechanisms that heterogeneous cell populations employ.

What I measured in this study are primarily *in vitro* cell culture models except for the human primary tumors. It is important to point out that each cell within heterogeneous populations of the breast cancer cells responds to Tamoxifen to different degrees and within such heterogeneous populations, complex response mechanisms help each cancer cell survive and adapt to the treatments. Our results from the generation of the Tamoxifen resistant derivatives show that different Tamoxifen resistant cell lines have different growth profiles and responses to therapeutic compounds. I predict that these models and their growth profiles are significantly simplified relative to what actually happens in patients' breast tumors undergoing Tamoxifen treatment. It is possible to speculate that $ER^+/HER2^-$ tumor cells may adapt to become Tamoxifen resistant similar to the way parental MCF7 cells become MCF7-TAM cells after long term Tamoxifen treatment, and $HER2^+/ER^+$ tumor cells may continue to grow in a similar manner as MCF7-HER2

cells with Tamoxifen treatment. In such situations, we could assume that heterogeneous cells that consist of $ER^+/HER2^-$ tumor cells and $HER2^+/ER^+$ tumor cells might utilize the common pathways that our *in vitro* cell line analyses have indicated, and therapeutic interventions, such as Dasatinib, may maximize the benefits from the endocrine treatment regimen. Our finding that dasatinib is effective against both models of the Tamoxifen resistance needs to be further tested before it reaches patients, particularly in higher models such as animal studies. The successful application of combinatorial treatment in patients requires that our finding must first be reproduced in mouse xenografts models and transgenic mouse models, which are more biologically relevant conditions.

Differences in molecular consequences of SERM and AI treatment

Modern endocrine therapy for breast cancer gives breast cancer patients chemotherapeutic treatment options other than SERMs such as Tamoxifen. The major alternatives are the Aromatase Inhibitors (AIs). While both SERMs and AIs target the estrogen receptor to inhibit tumor growth, the modes of inhibition are

fundamentally different, and it is very important to consider potential outcomes of each treatment regimen at a molecular scale. Tamoxifen directly inhibits binding of estrogen on ER to antagonize ER function, particularly at breast tumors. In contrast, it functions as an agonist in bones to maintain healthy bones which, for example, decrease risks of osteoporosis. This is beneficial especially for postmenopausal women who need endocrine replacement therapy. The negative aspects of the Tamoxifen and other SERMs are that they must modulate intracellular ER signaling pathways. While the consequences of changes in gene expressions are predicted, we do not fully understand the long-term outcomes of such modulations. One of the possible outcomes is an adaptation mechanism that cancer cells would use to rewire their signaling network in order to survive the pressure from the SERMs. Tamoxifen and other SERMs are not cytotoxic agents, rather they are cytostatic agents. While Tamoxifen and other SERMS such as raloxifen are favorable options for postmenapausal women who choose to benefit from the treatment for both breast cancer management and bone health improvement, it is important to consider benefits vs the detriment that rewiring of the signaling networks may cause in breast tumors with therapeutic resistance.

AIs may not be as appealing as SERMs to both premanosausal and postmenopausal women due to the lack of benefits for bone health, higher costs for treatment, and lack of large scale clinical trials and history compared to use of SERMs. The key difference between SERMs and AIs is that AIs do not directly interfere with ER in target tissues to inhibit the function of estrogen. AIs inhibit production of estrogen by inhibiting aromatase, a rate-limiting enzyme in estrogen biosynthesis, therefore limiting availability of estrogen for cancer cells. As there is no physical modulation of ER by AIs, ER signaling networks may be stopped rather than modulated in a case of SERM treatments. Which treatment would be better for long term effects at the molecular scale remains unclear at this point. Although AIs slow the growth of estrogen-dependent tumor cells resulting in clinical benefits, cancer cells are likely to adapt to estrogen-deprived conditions in order to survive such a condition. AI resistant breast tumors are emerging in breast cancer patients who were early adaptors of AI treatments. It is important to investigate resistance mechanisms to other endocrine treatments including AIs and other breast cancer treatments such as anti-HER2 treatment to benefit the entire spectrum of breast cancer patients.

In conclusion, phosphoproteomic analysis by mass spectrometry is a favorable platform to study therapeutic resistance problems because of its unique capability to collect data in an unbiased manner and facilitate discovery. Such analyses have important roles in revealing previously unknown signaling events from which biological hypotheses can be generated. Hypotheses can be tested in various assays to determine the roles of signaling events in controlling cellular phenotypes. In the future, phosphoproteomics will be more significant in the research and development of therapeutic compounds. The contributions that phosphoproteomics makes will be seen in many disease areas including therapeutic resistance to cancer treatments.

Acknowledgement

First of all, I want to emphasize that I would not be able to get this far in the graduate school without supports from many people. I would like to take a moment to sincerely thank all of them, and will mention their names in this acknowledgment. For me, spending time with you is the most precious treasure, and I cherish all of the memories I shared with you. I must apologize that I won't be able to mention all of your names here. Please let me say once more. Thank you very much!

I want to thank my thesis advisor and great mentor Forest White for welcoming me to the lab on September 2005 and guided me throughout the graduate school. Having an advisor like Forest is a true blessing and I feel very fortunate to receive various trainings under his supervision. I learned tremendously from interacting with Forest. I cannot imagine having a better advisor than Forest, and it was truly my pleasure working with him through out the graduate school.

I want to specially thank my excellent thesis committee members Doug Lauffenburger and Steve Tannenbaum. Doug is probably the person who I respect

the most. Doug truly cares about each student in the Department of Biological Engineering. As a chair of the thesis committee, his guidance was the most valuable for completion of the thesis. It was great experience for me to visit Astra Zeneca with Doug in July 2006. Steve is probably the nicest professor I know in MIT. Serving as a teaching assistant for Steve's class was an important and exciting experience for me. I was very happy that Steve cared so much about my thesis projects. I felt very lucky to have both of them in my thesis committee because I learned so much from them about many important things including, but not limited to science.

I want to thank the former and current members of the Forest White Lab. People in the White lab are very dynamic and active individuals who are exciting people to work with. Yi Zhang, Katrin Schmelzel, and Ji-Eun Kim were the three post-docs who were most senior members of the lab when I joined. They taught me the basics of capillary chromatography and running mass spectrometer for phosphoproteomic analysis. Former graduate students, Ale Wolf-Yadlin and Paul Huang are very good friend whose presence made the lab significantly more enjoyable. Presence of Ale in the mass spec lab made it so much fun to work

there. Paul really was the most productive graduate student I have met, and his knowledge and amount of efforts put in research blew my mind. Former post-doc Aleksandra Nita-Lazar is a good friend who is a good listener and helped me make progress on the thesis at various points. My collaborator Kartikeya Pant, or KP, from Tannenbaum was a very nice man who was fun to work with. I had two talented undergraduate students who worked with me. Rachel Niehus was a cheerful and active student who I felt very lucky to work with. Yi Wang is a disciplined and dedicated student. Both of them contributed tremendously to the experimental parts of the thesis. Current graduate students in the lab, Josh Apgar, Abhinav Arneja, Scott Carlson, Emily Miraldi, Bryan Owens, and Hui Liu are really nice and talented friends, and I thank all of them for all of their supports!

There are many friends outside the lab who I had lots of fun together both scientifically and non-scientifically. Shinya Umeno and Namiko Yamamoto are the Ph.D candidates in EECS and Aero&Astro at MIT who I shared the entire time while I was at MIT. We were in the Ph.D programs in different programs, and sharing ideas and thoughts made the graduate school so much more fun. Shin Kida and Yoshiaki Kuwata are the former Ph.D student in Marine Enginnering and

Aero&Astro at MIT whose guidance and supports were essential for me to make it to the end.

There are close friends who I have known since undergrad at UCLA and lived in Boston the part or entire time I was here. Calvin Jan is a labmate from UCLA, and we both got in graduate schools at several locations and ended up coming to Boston coincidentally. I consider him as a most reliable, humble friend who I trust the most. The 7+ years we spent together in LA and Boston were simply the best; no other word can describe how much I appreciate his presence. Another good friend from UCLA who moved to Boston is Aaron Chang. Both of Calvin and Aaron have the warm hearts that sincerely care for others. I very much feel fortunate to have good friends like you guys who stayed close to me.

I want to specially thank Katherine Schlieper for being very supportive and patient with me during the most difficult time of graduate school. The time we spent together is the most precious memory in the past nine years I spent in the USA.

Last, but not least, I want to thank my family. My younger sister Maria has been in the UK for over 10 years. She moved to UK right before I moved to USA

and we both have been studied abroad in a two different countries for a long time. The presence of younger sister who enjoy living abroad, and successfully finishing Medical degree from Cambridge University and Oxford University was a very positive influence for me. My grandparents Masaru and Yoshi have been the biggest supporter of me. I am well aware that they are proud of me and I hope that they enjoy seeing what I will do in the future. My stepfather Karl Benz has been very nice to everyone in the family, and his presence was huge. I want to show deepest appreciation to my mother Ekuko. She is a dependable person and a strong mother who is bold enough to happily study-abroad both of her kids when I and Maria were still teenagers. I have always felt that my mother will support me no matter when it is or where she is, and that support was essential for me to make it to the end of the grad school. I want to also thank my father Nicholai who passed away when I was four years old. He was a physicist from Moscow State University and he must be proud of me to finish the graduate school from MIT. This thesis is dedicated for him.

INDEX OF FIGURES AND TABLES

Chapter II.

Table 1

GI Number	Protein Name	Peptide Sequence	pY site	Parental MCF7				MCF7-HER2				MCF7-TAM			
				Control		Tamoxifen		Control		Tamoxifen		Control		Tamoxifen	
				Ratio	SD	Ratio	SD	Ratio	SD	Ratio	SD	Ratio	SD	Ratio	SD
gi|62990121	actin-like protein	HQGMMEGMHQKESpYVGK	Y53	1.05	-	1.67	-	1.00	-	0.80	-	--	-	--	-
gi|55743098	alpha 3 type VI collagen isoform 1 precursor	GERGFPGpYPGPK	Y2318	0.65	-	0.85	-	1.00	-	1.15	-	--	-	--	-
gi|45505182	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, ca	DHLEGLpYAK	Y1000	0.57	-	1.16	-	1.00	0.00	3.66	0.35	1.02	-	3.75	-
gi|38683797	ankyrin repeat and sterile alpha motif domain containing 1	EEDEHPpYELLLTAETK	Y454	1.16	0.25	1.19	0.06	1.00	0.00	1.09	0.05	0.77	0.13	0.98	0.14
gi|44662836	breast cancer anti-estrogen resistance 1	HLLAPGPQDIpYDVPPVR	Y249	0.90	-	1.44	-	1.00	0.00	1.27	0.29	1.06	0.24	1.38	0.01
		DVPDGPLLREETpYDVPPAFAK	Y327	1.22	0.22	0.87	0.06	1.00	0.00	1.22	0.20	0.68	0.26	0.79	0.09
		AQQGLpYQVPGPSPQFQSPPAK	Y128	1.12	-	1.09	-	1.00	0.00	1.05	0.12	0.61	-	0.47	-
		RPGPGTLpYDVPR	Y387	1.22	-	1.82	-	1.00	-	1.32	-	--	-	--	-
		GPNGRDPLLEVpYDVPPSVEK	Y287	0.97	-	1.18	-	1.00	0.00	1.45	0.24	0.90	-	1.67	-
		VLPPEVADGGVVDSGVpYAVPPPAER	Y410	--	-	--	-	1.00	-	0.66	-	0.37	-	0.55	-
gi|4503129	catenin, alpha-like 1	ALKLpTGVEGNLEALAEpYACK	T424/Y455	1.20	-	2.20	-	1.00	-	1.82	-	--	-	--	-
gi|10835010	catenin (cadherin-associated protein), delta 1	HYEDGYPGGSDNpYGSLSR	Y228	0.89	-	0.74	-	1.00	0.00	2.53	0.37	0.84	-	0.93	-
		SDNNpYSTPNER	Y904	0.95	-	1.01	-	1.00	0.00	2.54	0.39	1.16	0.32	1.17	0.07
gi|11034811	catenin (cadherin-associated protein), delta 2	ALQSPEHHIDPIpYEDR	Y424	1.03	0.14	1.31	0.03	1.00	0.00	3.14	0.43	0.53	-	0.90	-
		ASYAAGPASNpYADPYR	Y499	1.24	0.29	1.53	0.29	1.00	0.00	2.85	0.40	0.78	0.11	1.10	0.25
gi|4502709	cell division cycle 2 protein isoform 1	IGEGTpYGVVYK	Y15	0.81	0.15	0.25	0.03	1.00	0.00	0.43	0.11	0.49	0.12	0.34	0.06
gi|5031635	cofilin 1 (non-muscle)	HELQANCpYEEVKDR	Y139	--	-	--	-	1.00	0.00	1.45	0.03	0.59	0.05	0.80	0.10
gi|20357552	cortactin isoform a	LPSSPVpYEDAASFK	Y421	0.93	0.30	0.71	0.29	1.00	0.00	1.51	0.11	0.87	0.27	0.96	0.27
		GPVSGTEPEPVpYSMEAADYR	Y446	1.45	-	0.80	-	1.00	0.00	1.64	0.28	1.07	0.21	0.87	0.20
gi|4503049	cysteine-rich protein 2	GVNTGAVGSYIpYDRDPEGK	Y198	1.30	-	0.88	-	1.00	0.00	1.54	0.39	0.51	0.12	0.40	0.00
gi|18765758	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase	IYQpYIQSR	Y321	1.16	0.02	1.02	0.12	1.00	0.00	1.23	0.20	--	-	--	-
gi|4758250	ephrin B2	TADSVFCPHpYEK	Y304	--	-	--	-	1.00	-	5.83	-	2.48	-	3.91	-
gi|56119207	ephrin receptor EphA1	LWLJPpYVDLQAYEDPAQGALDFTR	Y599	0.44	-	0.74	-	1.00	-	2.47	-	--	-	--	-
gi|32967311	ephrin receptor EphA2	VLEDDPEATpYTTSGGK	Y772	2.60	0.29	2.03	0.19	1.00	0.00	1.93	0.28	2.16	0.31	2.28	0.20
		QSPEDVpYFSK	Y575	2.07	0.36	2.02	0.18	1.00	0.00	1.76	0.29	1.91	0.31	2.38	0.02
		TpYVDPHTpYEDPNQAVLK	Y588/Y594	3.84	-	3.24	-	1.00	0.00	3.13	0.35	2.59	0.24	2.58	0.08
		TYVDPHTpYEDPNQAVLK	Y594	2.03	0.20	1.53	0.42	1.00	0.00	2.27	0.41	1.92	0.35	2.10	0.01
		TpYVDPHTYEDPNQAVLK	Y588	2.48	0.38	1.55	0.31	1.00	0.00	1.59	0.25	1.47	0.32	1.72	0.18
gi|21361241	ephrin receptor EphA3 isoform a precursor	VLEDDPEAApYTTR	Y779/Y779/Y833	1.14	0.20	1.09	0.07	1.00	0.00	1.71	0.33	0.95	0.16	1.19	0.20
		TpYVDPHTpYEDPTQAVHEFAK	Y596/Y602	0.22	0.02	0.21	0.15	1.00	0.00	3.42	0.32	0.17	0.15	0.18	0.16
		TpYVDPHTYEDPTQAVHEFAK	Y596	0.27	-	0.10	-	1.00	0.00	1.85	0.24	0.12	-	0.12	-
gi|4758280	ephrin receptor EphA4	TYVDPFTpYEDPNQAVR	Y602	1.55	0.23	0.61	0.13	1.00	0.00	1.56	0.22	0.54	0.19	1.48	0.20
		TpYVDPFTYEDPNQAVR	Y596	1.19	0.17	0.60	0.01	1.00	0.00	1.87	0.13	0.83	0.12	1.55	0.32
gi|4758282	ephrin receptor EphA7	VIEDDPEAVpYTTTGGK	Y791	1.49	0.14	1.02	0.16	1.00	0.00	2.91	0.40	0.75	-	1.14	-
		TYIDPETpYEDPNR	Y614	2.56	-	1.87	-	1.00	0.00	3.94	0.15	0.70	-	1.21	-
		ADQEGDEELpYFHFK	Y597	1.68	0.15	0.99	0.23	1.00	0.00	3.13	0.33	0.95	0.07	1.11	0.17
gi|32528301	ephrin receptor EphB4 precursor	FLEENSSDPTpYTSSLGGK	Y774	3.55	-	1.47	-	1.00	0.00	3.20	0.22	1.82	-	2.23	-
		EAEpYSDKHGQpYLIGHGTK	Y574/Y581	2.83	-	3.14	-	1.00	-	7.63	-	--	-	--	-
gi|54792096	erbB-2 isoform a	LLDIDETEpYHADGGKVPIK	Y877	0.10	0.03	0.02	0.01	1.00	0.00	0.53	0.12	0.10	0.01	0.09	0.03
		GTPTAENPEpYLGLDVPV	Y1248	0.03	-	0.03	-	1.00	0.00	2.38	0.34	0.03	-	0.02	-
		YSEDPTVPLPSETDGpYVAPLTCSPQPEYVNQPDVRPQPPSPR	Y1127	0.03	-	0.51	-	1.00	0.00	1.09	0.34	0.03	0.02	0.56	0.19
		FVVIQNEDLGPASPLDSTFpYR	Y1005	0.23	-	0.29	-	1.00	-	3.38	-	--	-	--	-
		GAPPSpTFJGTPTAENPEpYLGLDVPV	T1236/Y1248	--	-	--	-	1.00	-	2.65	-	0.03	-	0.02	-
gi|8923909	ERBB2 interacting protein isoform 2	RAQIPEGDpYLSYR	Y1104	0.87	0.04	0.68	0.04	1.00	0.00	2.20	0.35	0.71	0.19	0.67	0.17
gi|54792100	erbB-3 isoform 1 precursor	SLEATDSAFDNPDpYWHSR	Y1328	0.30	0.09	0.19	0.11	1.00	0.00	0.60	0.08	0.29	-	0.19	-
gi|4503471	eukaryotic translation elongation factor 1 alpha 1	EHALLApYTLGVK	Y141	1.01	-	0.59	-	1.00	0.00	0.68	0.24	0.53	0.13	0.58	0.07
gi|4503787	fyn-related kinase	HGHpYFVALFDYQAR	Y46	3.17	-	1.66	-	1.00	0.00	1.51	0.12	2.95	-	0.56	-
gi|88861553	glucocorticoid receptor DNA binding factor 1	NEEENIpYSVPHDSTQGK	Y1105	0.65	0.31	0.48	0.19	1.00	0.00	1.31	0.39	0.69	0.22	0.89	0.28
gi|21614520	glucose-6-phosphate dehydrogenase	VQPNEAVpYTK	Y401	--	-	--	-	1.00	-	0.55	-	0.21	-	0.30	-
		VGFQpYEGTYK	Y503	--	-	--	-	1.00	-	0.57	-	0.45	-	0.31	-
gi|49574532	glycogen synthase kinase 3 alpha	GEPNVSpYICSR	Y279	1.08	0.02	0.97	0.20	1.00	0.00	1.02	0.33	1.03	0.24	1.03	0.31
gi|41393573	G protein-coupled receptor kinase interactor 1	LQPFHSTELEDDAIpYSVHVPAGLYR	Y545	1.18	0.06	1.11	0.30	1.00	0.00	1.16	0.16	0.80	-	0.84	-
gi|46370071	GRB2-associated binding protein 1 isoform a	DASSQDCpYDIPR	Y406	0.88	-	0.15	-	1.00	-	0.93	-	--	-	--	-
gi|38201640	homeodomain-interacting protein kinase 1 isoform 1	AVCSTpYLQSR	Y352	0.94	0.07	0.45	0.03	1.00	0.00	0.57	0.16	0.80	-	0.73	-
gi|45827723	hypothetical protein LOC25895 isoform b	GSSVRGGCpYHH	Y137	1.94	-	0.39	-	1.00	-	1.23	-	--	-	--	-
gi|12232415	hypothetical protein LOC64762	SELPpYEELWLEEGKPSHQPLTR	Y453	--	-	--	-	1.00	0.00	2.55	0.15	0.53	0.10	0.55	0.08
gi|50843820	hypothetical protein LOC56243	NEGFpYADPYLYHEGR	Y393	0.82	-	1.29	-	1.00	0.00	1.37	0.35	0.74	0.18	0.74	0.02
gi|4755142	inositol polyphosphate phosphatase-like 1	LpYEWISIDKDEAGAK	Y886	--	-	--	-	1.00	0.00	1.58	0.25	0.95	0.32	1.03	0.17
gi|22325385	intersectin 2 isoform 1	REEPEALpYAAVNK	Y967	0.75	0.11	0.80	0.04	1.00	0.00	1.38	0.25	0.34	0.04	0.34	0.08
gi|4826776	Janus kinase 2	EVGDpYGQLHETEVLLK	Y570	1.04	-	0.54	-	1.00	-	0.58	-	--	-	--	-
gi|4504919	keratin 8	AQpYEDIANR	Y267	6.65	-	8.83	-	1.00	0.00	1.14	0.07	0.57	-	0.40	-
gi|4557888	keratin 18	DWSHZFJ	Y128	1.28	-	0.36	-	1.00	-	0.16	-	--	-	--	-
gi|24234699	keratin 19	SLLEGQEDHpYNNLSASK	Y391	--	-	--	-	1.00	-	1.33	-	1.14	-	1.24	-
gi|31559819	keratin 25C	TDLEIQLETLSEELApYLKK	Y282	0.64	-	0.08	-	1.00	-	0.11	-	--	-	--	-
gi|54607120	lactotransferrin	YpYGYTGAFR	Y547	0.90	0.33	2.31	0.40	1.00	0.00	1.82	0.29	0.72	0.20	0.94	0.11
gi|45505163	LISCH protein isoform 1	SRDPHpYDDFR	Y551	1.35	0.07	0.65	0.12	1.00	0.00	0.61	0.06	1.37	-	1.56	-
		VLpYYMEJ	Y406	--	-	--	-	1.00	0.00	1.00	0.08	1.76	0.02	1.10	0.07
gi|66932916	mitogen-activated protein kinase 1	VADPDHDHTGFLTEpYVATR	Y186	0.47	0.09	0.16	0.04	1.00	0.00	0.65	0.12	0.44	0.01	0.42	0.05
gi|91718899	mitogen-activated protein kinase 3 isoform 1	IADPEHDHTGFLTEpYVATR	Y204	0.49	0.05	0.20	0.04	1.00	0.00	0.55	0.06	0.50	0.09	0.47	0.10
gi|20986516	mitogen-activated protein kinase 14 isoform 4	HTDDEMTGpYVATR	Y181	0.92	-	0.29	-	1.00	-	0.64	-	--	-	--	-
gi|10047086	mitogen-inducible gene 6 protein	VSSTHpYYLLPERPPYLDK	Y394	0.36	-	0.23	-	1.00	0.00	0.57	0.11	0.16	0.00	0.21	0.08
gi|60685216	muscle-derived protein 77	KpSpTIDQpYELR	S284/T285/Y289	0.25	-	4.95	-	1.00	-	4.35	-	--	-	--	-
gi|4506357	myelin protein zero-like 1 isoform a	SESVVpYADIR	Y263	0.63	-	0.60	-	1.00	0.00	0.48	0.21	0.19	-	0.57	-
gi|21361181	Na+/K+ -ATPase alpha 1 subunit isoform a proprotein	GIVVpYTGDR	Y260	1.09	0.14	1.96	0.14	1.00	0.00	1.75	0.25	1.08	0.09	1.85	0.12
gi|5453680	neural precursor cell expressed, developmentally down-regulated	EKDpYDFPPPMR	Y241	0.83	-	0.81	-	1.00	0.00	1.79	0.26	0.65	0.17	0.58	0.22
gi|4506345	paxillin	VGEEEHVpYSFPNK	Y118	1.22	0.14	1.50	0.30	1.00	0.00	2.14	0.36	0.53	0.17	0.90	0.02
		FIHQQPQSSSPVpYGSSAK	Y88	1.07	0.14	0.75	0.15	1.00	0.00	1.77	0.03	--	-	--	-
gi|31377667	peroxisomal LON protease-like	pYAVQLVEMLDMSVPAVAK	Y153	1.52	-	1.64	-	1.00	-	1.46	-	--	-	--	-
gi|4826908	phosphoinositide-3-kinase, regulatory subunit 2 (p85 beta)	NETEDQpYALMEDEDDLPHHEER	Y605	0.33	0.07	0.10	0.03	1.00	0.00	1.70	0.22	0.25	0.01	0.48	0.03
		SREYDQLpYEEYTR	Y464	--	-	--	-	1.00	-	1.08	-	0.54	-	0.60	-
gi|32455248	phosphoinositide-3-kinase, regulatory subunit, polypeptide 1 isofo	LNEWLGNENTEDQpYSLVEDDEDLPHHDEK	Y607	--	-	--	-	1.00	0.00	1.26	0.10	0.29	0.04	0.43	0.06
gi|21361831	partitioning-defective protein 3 homolog	ERDpYAEIQDFHR	Y1080	1.67	0.19	1.05	0.19	1.00	0.00	1.40	0.07	--	-	--	-

gi|37595548	phosphoinositol 3-phosphate-binding protein-3	SEDIpYADPAAYVMR	Y492	0.86	0.04	0.83	0.01	1.00	0.00	1.78	0.27	0.93	0.11	0.94	0.19
gi|6005830	plakophilin 3	NLIpYDNADNK	Y390	1.70	-	1.06	-	1.00	0.00	3.82	0.31	0.81	0.07	1.10	0.19
		ADpYDTLSLR	Y176	--	-	--	-	1.00	0.00	1.88	0.25	1.07	0.14	1.07	0.13
gi|53829374	plakophilin 4 isoform a	STTNpYVDFYSTK	Y1168	1.31	-	1.08	-	1.00	0.00	2.75	0.32	0.67	-	0.77	-
		LYLQSPHSpYEDPYFDDR	Y1139	1.19	0.61	0.70	0.00	1.00	0.00	3.80	0.49	0.85	0.15	1.15	0.10
		NNYALNTTATpYAEPYRPIQYR	Y478	1.53	0.10	1.64	0.19	1.00	0.00	3.83	0.54	0.98	0.00	1.11	0.11
		NNpYALNTTATpYAEPYRPIQYR	Y470/Y478	2.39	-	1.50	-	1.00	0.00	4.34	0.80	0.63	0.26	1.11	0.22
		TVHDMEQFGQQQYDIpYER	Y372	2.83	-	1.61	-	1.00	0.00	3.68	0.40	1.16	0.25	1.08	0.25
		VHFPASTDpYSTQYGLK	Y1156	2.03	-	1.35	-	1.00	-	7.28	-	--	-	--	-
gi|89028506	PREDICTED: hypothetical protein	EATQPEPIpYAESTK	Y413	1.10	0.00	0.61	0.14	1.00	0.00	0.52	0.07	0.92	0.04	0.88	0.08
gi|4503823	protein-tyrosine kinase fyn isoform a	LIEDNEpYTAR	Y420/Y394/Y418/Y425	0.51	-	0.57	-	1.00	0.00	1.73	0.22	0.55	0.21	0.58	0.19
gi|4506303	protein tyrosine phosphatase, receptor type, A isoform 1 precurs	VVQEYIDAFSDpYANFK	Y798	1.27	-	1.46	-	1.00	0.00	1.47	0.14	1.03	0.15	1.57	0.19
gi|33356177	protein tyrosine phosphatase, non-receptor type 11	IQNTGDpYYDLYGGEK	Y62	0.38	-	0.38	-	1.00	0.00	1.45	0.02	0.30	0.13	0.28	0.07
gi|24476013	PTK2 protein tyrosine kinase 2 isoform a	YMEDSTpYYK	Y576	0.76	0.23	0.63	0.12	1.00	0.00	1.54	0.36	0.61	-	0.79	-
gi|34147513	RAB7, member RAS oncogene family	QETEVELpYNEFPEPIKLDK	Y183	0.27	0.00	0.19	0.04	1.00	0.00	3.17	0.42	0.36	-	0.29	-
gi|26986534	Rho GTPase activating protein 12	ATTPPNQGRPDpSPVpYANLQELK	S240/Y243	0.96	-	0.64	-	1.00	-	1.07	-	--	-	--	-
gi|89276756	serine/threonine-protein kinase PRP4K	LCDFGSASHVADNDITPpYLVSR	Y849	1.03	0.08	0.60	0.08	1.00	0.00	0.83	0.13	0.94	0.07	0.97	0.11
gi|4506935	SHB (Src homology 2 domain containing) adaptor protein B	GESAGpYMEPYEAQR	Y268	1.18	0.32	1.27	0.07	1.00	0.00	1.73	0.22	1.03	0.26	1.37	0.14
		LDpYCGGSGEPGGVQR	Y114	1.28	0.19	1.45	0.04	1.00	0.00	2.56	0.43	1.00	0.19	1.26	0.31
		DKVTIADDpYSDPFDAK	Y246	1.13	0.09	1.14	0.29	1.00	0.00	2.39	0.46	1.10	0.29	1.10	0.02
		LPQDDDRPADEpYDQPWEWNR	Y336	0.97	0.22	0.82	0.04	1.00	0.00	1.85	0.41	0.85	0.10	0.95	0.16
		GIQLpYDTPYEPEGQSVDSDSESTVSPR	Y297	--	-	--	-	1.00	-	1.52	-	0.87	-	1.25	-
gi|32261324	SHC (Src homology 2 domain containing)	ELFDDPSpYVNVQNLDK	Y427	0.23	-	0.23	-	1.00	0.00	0.86	0.14	0.33	-	0.26	-
gi|45592957	SNAP25-interacting protein	GEGLpYADPYGLLHEGR	Y268	0.90	0.31	0.77	0.01	1.00	0.00	2.00	0.36	0.65	0.23	1.10	0.18
		JEPLpYAAFPGSHLTNGDLR	Y136	1.25	-	0.80	-	1.00	0.00	2.26	0.17	1.11	-	1.19	-
gi|4506903	splicing factor, arginine/serine-rich 9	GSPHpYFSPFRPY	Y214	0.69	0.19	0.44	0.05	1.00	0.00	0.64	0.09	0.55	-	0.66	-
gi|32401445	sprouty-related protein with EVH-1 domain 2	GEVPJHDpYNYPYVDSSDFGLGEDPK	Y264	0.15	-	0.09	-	1.00	-	0.34	-	--	-	--	-
		HDYNYPpYVDSSDFGLGEDPK	Y268	0.20	-	0.19	-	1.00	-	0.97	-	--	-	--	-
gi|10863921	synapse-associated protein 102	RDNEVDGQDpYHFVVSR	Y673	0.56	-	0.62	-	1.00	0.00	1.06	0.06	0.43	-	0.45	-
gi|56549666	tyrosine kinase, non-receptor, 2 isoform 1	ALPQNDDHpYVMQEHR	Y284	0.58	-	0.61	-	1.00	-	0.62	-	--	-	--	-
		KPTpYDPVSEDQDPLSSDFK	Y518	0.73	0.18	0.66	0.07	1.00	0.00	1.20	0.19	0.46	0.02	0.55	0.08
		VSSTHpYYLLPERPSYLER	Y857	0.77	0.14	0.52	0.06	1.00	0.00	0.97	0.15	0.44	-	0.61	-
gi|42518070	tight junction protein 2 (zona occludens 2) isoform 1	HPDIpYAVPIK	Y1118	0.88	0.01	0.94	0.04	1.00	0.00	1.13	0.19	0.89	0.00	1.22	0.00
gi|38787941	tensin like C1 domain containing phosphatase isoform 1	GPLDGSPpYAQVQRPPR	Y483	0.53	-	0.71	-	1.00	0.00	1.28	0.21	0.44	-	0.37	-
gi|31543838	tyrosine kinase 2	LLAQAEGEPCpYIR	Y292	1.50	-	1.41	-	1.00	0.00	1.43	0.17	0.75	0.03	1.58	0.33
		pYKGCCEDQGEK	Y962	0.15	-	0.14	-	1.00	-	0.16	-	--	-	--	-
gi|46430499	v-rel reticuloendotheliosis viral oncogene homolog A, nuclear fact	TPPZADPSLQAPVR	Y257	--	-	--	-	1.00	-	1.50	-	1.58	-	2.63	-
gi|40254887	5'-3' exoribonuclease 1	FEKpYLVK	Y339	0.57	-	1.16	-	1.00	-	3.91	-	--	-	--	-

Table 2: Twenty one pY sites detected in both anaylytical replicates in MCF7-TAM with NOS inhibitors treat

Numbers in grey row are averages, and numbers in white row are standard errors

Protein Name	Sequence	Control	1400W	SMTCH	SMTCL
breast cancer anti-estrogen resist	AQQGLpYQVPGPSPQFQSPF	1.00	1.19	1.74	1.40
		0.00	0.39	0.59	0.53
breast cancer anti-estrogen resist	DVPDGPLLREETpYDVPPAFA	1.00	0.98	1.24	0.99
		0.00	0.04	0.09	0.13
breast cancer anti-estrogen resist	HLLAPGPQDIpYDVPPVR	1.00	1.05	1.01	0.86
		0.00	0.09	0.26	0.13
glycogen synthase kinase 3 beta	GEPNVSpYICSR	1.00	0.82	0.78	0.74
		0.00	0.05	0.06	0.07
ephrin receptor EphA2	BEQLKPLKTpYVDPHTYEDPN	1.00	0.77	2.40	1.99
		0.00	0.01	0.05	0.18
ephrin receptor EphA2	TpYVDPHTYEDPNQAVLK	1.00	0.94	0.92	0.73
		0.00	0.09	0.11	0.16
ephrin receptor EphA2	QSPEDVpYFSK	1.00	1.26	1.68	0.99
		0.00	0.00	0.25	0.07
ephrin receptor EphA2	VLEDDPEATpYTTSGGK	1.00	1.23	1.94	1.42
		0.00	0.05	0.07	0.27
paxillin	VGEEEHVYpSFPNK	1.00	1.47	1.59	1.34
		0.00	0.13	0.27	0.17
cyclin-dependent kinase 2	IGEGTpYGVVUK	1.00	0.80	0.46	0.79
		0.00	0.05	0.01	0.09
serine/threonine-protein kinase P	LCDFGSASHVADNDITPpYLV	1.00	0.80	0.79	0.68
		0.00	0.00	0.06	0.10
mitogen-activated protein kinase 3	IADPEHDHTGFLTEpYVATR	1.00	0.66	0.69	0.41
		0.00	0.07	0.05	0.12
mitogen-activated protein kinase 1	VADPDHDHTGFLTEpYVATR	1.00	0.65	0.74	0.41
		0.00	0.05	0.13	0.05
myelin transcription factor 1-like	RpYPGEVTIPTPKPK	1.00	3.79	3.64	2.93
		0.00	2.05	1.75	1.17
intersectin 2 isoform 1 ; intersecti	REEPEALpYAAVNK	1.00	1.39	1.67	2.06
		0.00	0.42	0.39	0.86
ERBB2 interacting protein isoform	AQIPEGDpYLSYR	1.00	0.92	1.07	0.87
		0.00	0.14	0.26	0.10
PREDICTED: hypothetical protein	EATQPEPIpYAESTK	1.00	0.94	1.54	1.37
		0.00	0.33	0.67	0.49
tyrosine kinase, non-receptor, 2 is	KPTpYDPVSEDQDPLSSDFKF	1.00	0.96	0.96	1.01
		0.00	0.08	0.19	0.05
SHB (Src homology 2 domain cont	AGKGESAGpYMEPYEAQR	1.00	1.23	2.20	1.85
		0.00	0.21	1.51	1.33
SHB (Src homology 2 domain cont	VTIADDpYSDPFDAKNDLK	1.00	1.01	2.56	1.94
		0.00	0.09	0.36	0.38
ankyrin repeat domain 3	VGSGGFGQVpYK	1.00	0.82	0.55	0.75
		0.00	0.11	0.06	0.11

Table 3: Twenty two pY sites detected in both anaylytical replicates in MCF7-HER2 with NOS inhibitors trea

Numbers in grey row are averages, and numbers in white row are standard errors

Protein Name	Sequence	Control	1400W	SMTCH	SMTCL
SHB (Src homology 2 domain conta	AGKGESAGpYMEPYEAQR	1.00	1.31	1.44	1.59
		0.00	0.09	0.38	0.15
SHB (Src homology 2 domain conta	DKVTIADDpYSDPFDAK	1.00	1.35	1.71	1.46
		0.00	0.75	1.00	1.60
SHB (Src homology 2 domain conta	LDpYCGGSSKGEPGGVQR	1.00	1.51	1.41	2.33
		0.00	0.02	0.25	0.12
paxillin	VGEEEHVpYSFPNK	1.00	1.96	1.43	1.52
		0.00	0.20	0.11	0.09
catenin (cadherin-associated protei	ALQSPEHHIDPIpYEDR	1.00	1.63	2.77	2.15
		0.00	0.41	0.46	0.61
catenin (cadherin-associated protei	ASYAAGPASNpYADPYR	1.00	1.21	1.03	1.25
		0.00	0.11	0.12	0.16
ephrin receptor Eph4 or EphA5	VLEDDPEAApYTTR	1.00	1.60	1.69	1.26
		0.00	0.27	0.04	0.23
ERBB2 interacting protein isoform 2	AQIPEGDpYLSYR	1.00	1.33	1.23	1.29
		0.00	0.14	0.13	0.52
glycogen synthase kinase 3 beta	GEPNVSpYICSR	1.00	1.23	1.08	1.23
		0.00	0.09	0.26	0.19
tyrosine kinase, non-receptor, 2 isof	ALPQNDDHpYVMQEHR	1.00	1.80	1.46	1.49
		0.00	0.28	0.58	0.15
tyrosine kinase, non-receptor, 2 isof	KPTpYDPVSEDQDPLSSDFKF	1.00	1.38	1.09	1.00
		0.00	0.09	0.36	0.14
partitioning-defective protein 3 hom	ERDpYAEIQDFHR	1.00	1.29	1.25	0.93
		0.00	0.58	0.03	0.24
intersectin 2 isoform 3 ; intersectin 2	REEPEALpYAAVNK	1.00	2.38	1.43	2.29
		0.00	0.81	1.01	1.68
ephrin receptor EphA7	VIEDDPEAVpYTTTGGK	1.00	1.67	1.59	1.76
		0.00	0.54	0.74	0.65
tensin like C1 domain containing ph	GPLDGSPpYAQVQRPPR	1.00	0.99	0.73	1.24
		0.00	0.09	0.02	0.25
mitogen-activated protein kinase 3 i	IADPEHDHTGFLTEpYVATR	1.00	1.28	0.62	1.41
		0.00	0.18	0.09	0.25
mitogen-activated protein kinase 1 ;	VADPDHDHTGFLTEpYVATR	1.00	1.45	0.68	1.31
		0.00	0.09	0.03	0.09
phosphoinositide-3-kinase, regulato	NETEDQpYALMEDEDDLPHHI	1.00	1.79	1.68	1.67
		0.00	0.02	0.09	0.13
keratin 19	SLLEGQEDHpYNNLSASK	1.00	1.05	0.88	1.10
		0.00	0.02	0.34	0.00
cell division cycle 2 protein isoform	IGEGTpYGVVYK	1.00	1.09	0.65	1.23
		0.00	0.08	0.02	0.19
homeodomain-interacting protein ki	AVCSTpYLQSR	1.00	1.07	3.02	1.30
		0.00	0.13	0.65	0.14
phosphorylase kinase, alpha 1 (mus	ODLSpYLCSR	1.00	1.19	1.20	0.49
		0.00	0.24	0.29	0.17

Table 4: Phosphoproteomic analysis of EML4-ALK signaling events

Abundance relative to the highest iTRAQ peak area

#	Acronym	KM(%)	ALK(%)	V1(%)	V3b(%)
1	ALK Y1096	3	61	57	100
2	EML4 Y125	6	9	66	100
3	EML1 Y184	2	2	69	100
4	EML1 Y396	44	52	100	74
5	HSP90 Y483	7	35	100	60
6	SDN Y109	4	45	69	100
7	RAN Y147	22	52	100	41
8	RAN Y155	6	36	100	17
9	PKM2 Y390	4	59	100	55
10	PKM2 Y105	16	36	100	67
11	PKM2 Y175	18	61	100	51
12	PGK1 Y196	9	37	100	33
13	LDH-A Y238	3	100	80	55
14	LPP Y302	100	57	34	35
15	FER Y715	21	100	42	69
16	GRF-1 Y1106	92	95	89	76
17	ERBIN Y1097	100	64	30	29
18	PARD3 Y1076	98	100	55	61
19	PTTG1IP Y171	4	100	9	3
20	GAB1 Y407	70	65	76	100
21	IRS-2 Y734	13	13	29	100
22	CTNND1 Y904	100	53	34	43
23	CAS-L Y165	82	100	87	35
24	PXN Y118	100	69	39	45
25	TENSIN 1 Y1480	100	43	15	16
26	HIPK1 Y352	99	100	93	97
27	PROFILIN 1 Y128	29	74	100	43
28	HIP14 Y105	3	100	4	7
29	HIPK3 Y359	79	100	82	60
30	AP2B1 Y276	6	100	14	12
31	ACLY Y672	57	100	78	63
32	TENSIN 2 Y483	100	74	43	50
33	CAVEOLIN-1 Y14	100	37	23	25
34	CDC2 Y15	77	47	89	89
35	DYRK1 Y321	96	96	86	95
36	Eph A3 Y779	100	15	14	18
37	FYN Y420	55	100	71	77
38	GMD Y323	9	100	45	51
39	GSK3 Y216	94	97	93	90
40	ERK1 Y205	42	100	52	69
41	p38-beta Y182	36	100	50	69
42	P38 ALPHA Y181	100	61	44	62
43	KIRREL Y622	100	90	69	78
44	PSAT1 Y346	32	90	96	58
45	SHP-2 Y584	36	75	61	100
46	SHP-2 Y62	53	49	62	100
47	PRP4 Y849	85	96	91	98
48	MAGI-1 Y373	67	100	77	73
49	PRP0 Y24	33	54	100	54
50	VIGILIN Y437	25	100	47	67
51	TALIN 1 Y70	58	100	44	34
52	FAK Y576	100	66	47	53
53	ANXA2 Y23	30	100	39	32
54	ANXA2 Y315	32	94	92	70
55	ANXA2 Y234	48	100	53	72
56	ANXA5 Y93	61	100	85	60
57	PIK3R1 Y580	52	57	76	100
58	PIK3R2 Y458	62	100	54	49
59	ENO Y24	35	64	84	100
60	ENO Y43	30	100	48	57
61	SFRS9 Y215	62	89	70	100
62	PSMA2 Y56	25	100	90	66
63	DOK1 Y450	89	100	69	92
64	N-WASP Y253	100	84	66	64
65	ACTIN Y220	7	40	37	100
66	DBI Y77	13	100	31	25
67	ELMO2 Y48	60	88	77	98
68	SGK269 Y632	93	99	81	77
69	NCK1 Y105	61	100	67	67
70	NCK2 Y110	24	100	82	63
71	PTPRA Y825	100	86	65	81
72	RPS10 Y12	100	76	59	40
73	SHC Y423	19	30	43	100
74	LOC328092 Y162	10	100	33	44
75	GCP2 Y83	43	62	85	100
76	ASS Y236	51	100	83	61
77	WDR Y237	34	81	100	63
78	FRS2 Y348	60	100	58	79
79	ITSN Y921	23	100	58	48
80	RACK1 Y228	42	43	100	53
81	eIF3-eta Y514	30	62	73	100
82	LDH-B Y240	19	100	57	69
83	VCP Y644	36	79	71	100
84	ATP1A1 Y260	69	100	43	50

Fold change relative to Kinase Mutant

Acronym	KM/KM	ALK/KM	V1/KM	V3b/KM
ALK Y1096	1.0	19.9	18.6	32.5
EML4 Y125	1.0	1.5	11.9	18.0
EML1 Y184	1.0	1.1	35.2	51.2
EML1 Y396	1.0	1.2	2.3	1.7
HSP90 Y483	1.0	4.9	14.0	8.4
SDN Y109	1.0	11.7	18.2	26.3
RAN Y147	1.0	2.3	4.5	1.8
RAN Y155	1.0	5.8	16.2	2.8
PKM2 Y390	1.0	13.2	22.3	12.4
PKM2 Y105	1.0	2.2	6.1	4.1
PKM2 Y175	1.0	3.4	5.5	2.8
PGK1 Y196	1.0	4.3	11.7	3.8
LDH-A Y238	1.0	33.2	26.6	18.2
LPP Y302	1.0	0.6	0.3	0.4
FER Y715	1.0	4.7	2.0	3.3
GRF-1 Y1106	1.0	1.0	1.0	0.8
ERBIN Y1097	1.0	0.6	0.3	0.3
PARD3 Y1076	1.0	1.0	0.6	0.6
PTTG1IP Y171	1.0	23.8	2.1	0.8
GAB1 Y407	1.0	0.9	1.1	1.4
IRS-2 Y734	1.0	1.0	2.1	7.5
CTNND1 Y904	1.0	0.5	0.3	0.4
CAS-L Y165	1.0	1.2	1.1	0.4
PXN Y118	1.0	0.7	0.4	0.4
TENSIN 1 Y1480	1.0	0.4	0.2	0.2
HIPK1 Y352	1.0	1.0	0.9	1.0
PROFILIN 1 Y128	1.0	2.5	3.4	1.5
HIP14 Y105	1.0	32.7	1.4	2.5
HIPK3 Y359	1.0	1.3	1.0	0.8
AP2B1 Y276	1.0	17.6	2.5	2.2
ACLY Y672	1.0	1.7	1.4	1.1
TENSIN 2 Y483	1.0	0.7	0.4	0.5
CAVEOLIN-1 Y14	1.0	0.4	0.2	0.2
CDC2 Y15	1.0	0.6	1.2	1.2
DYRK1 Y321	1.0	1.0	0.9	1.0
Eph A3 Y779	1.0	0.2	0.1	0.2
FYN Y420	1.0	1.8	1.3	1.4
GMD Y323	1.0	11.7	5.3	6.0
GSK3 Y216	1.0	1.0	1.0	1.0
ERK1 Y205	1.0	2.4	1.2	1.7
p38-beta Y182	1.0	2.8	1.4	1.9
P38 ALPHA Y181	1.0	0.6	0.4	0.6
KIRREL Y622	1.0	0.9	0.7	0.8
PSAT1 Y346	1.0	2.8	3.0	1.8
SHP-2 Y584	1.0	2.1	1.7	2.7
SHP-2 Y62	1.0	0.9	1.2	1.9
PRP4 Y849	1.0	1.1	1.1	1.2
MAGI-1 Y373	1.0	1.5	1.1	1.1
PRP0 Y24	1.0	1.6	3.0	1.6
VIGILIN Y437	1.0	4.0	1.9	2.7
TALIN 1 Y70	1.0	1.7	0.7	0.6
FAK Y576	1.0	0.7	0.5	0.5
ANXA2 Y23	1.0	3.4	1.3	1.1
ANXA2 Y315	1.0	3.0	2.9	2.2
ANXA2 Y234	1.0	2.1	1.1	1.5
ANXA5 Y93	1.0	1.6	1.4	1.0
PIK3R1 Y580	1.0	1.1	1.5	1.9
PIK3R2 Y458	1.0	1.6	0.9	0.8
ENO Y24	1.0	1.8	2.4	2.9
ENO Y43	1.0	3.4	1.6	1.9
SFRS9 Y215	1.0	1.4	1.1	1.6
PSMA2 Y56	1.0	3.9	3.5	2.6
DOK1 Y450	1.0	1.1	0.8	1.0
N-WASP Y253	1.0	0.8	0.7	0.6
ACTIN Y220	1.0	5.7	5.2	14.2
DBI Y77	1.0	7.7	2.4	1.9
ELMO2 Y48	1.0	1.5	1.3	1.6
SGK269 Y632	1.0	1.1	0.9	0.8
NCK1 Y105	1.0	1.6	1.1	1.1
NCK2 Y110	1.0	4.2	3.5	2.7
PTPRA Y825	1.0	0.9	0.7	0.8
RPS10 Y12	1.0	0.8	0.6	0.4
SHC Y423	1.0	1.5	2.2	5.1
LOC328092 Y162	1.0	10.4	3.4	4.6
GCP2 Y83	1.0	1.4	2.0	2.3
ASS Y236	1.0	1.9	1.6	1.2
WDR Y237	1.0	2.4	2.9	1.9
FRS2 Y348	1.0	1.7	1.0	1.3
ITSN Y921	1.0	4.3	2.5	2.1
RACK1 Y228	1.0	1.0	2.4	1.2
eIF3-eta Y514	1.0	2.1	2.5	3.4
LDH-B Y240	1.0	5.3	3.0	3.7
VCP Y644	1.0	2.2	2.0	2.8
ATP1A1 Y260	1.0	1.5	0.6	0.7